The *trfA* and *trfB* promoter regions of broad host range plasmid RK2 share common potential regulatory sequences

Christopher A.Smith, Victoria Shingler and Christopher M.Thomas*

Department of Genetics, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, UK

Received 21 February 1984; Revised and Accepted 21 March 1984

ABSTRACT

The positions of the trfA and trfB promoters of broad host range IncP plasmid RK2 (identical to RP1, RP4, R68 and R18) were identified by RNA polymerase protection studies, and the nucleotide sequences of the promoter regions determined. A mutation within the trfA promoter sequence is associated with loss of kilD activity. In addition a probable promoter region for the kilB locus was identified. The three promoter regions share common palindromic sequences which may serve as sites for the coordinate regulation of replication and kil functions.

INTRODUCTION

RK2 is one member of a group of indistinguishable plasmid isolates which are able to transfer themselves between and to maintain themselves in diverse Gram-negative bacterial species (1,2,3,4,5). These isolates belong to Escherichia coli incompatibility group P and include RP1, RP4, R68 (6) and R18 (7) as well as RK2. Fig. 1 shows a general genetic map of RK2, which carries genes conferring resistance to the antibiotics tetracycline, kanamycin and penicillin (the latter being carried on the transposon Tn1) as well as genes involved in conjugal transfer which occupy about one third of its genome. Vegetative replication of RK2 proceeds unidirectionally from a unique origin, $oriV_{RK2}$ (15), and requires the product of a trans-acting gene trfA (16) separated from $oriV_{RK2}$ by the tetracycline resistance determinant. It has been shown that these two loci are sufficient to form a mini-RK2 replicon functional at least in E.coli (16), Pseudomonas putida (17) and P.aeruginosa (A.A.K. Hussain & C.M. Thomas, unpublished). However, in addition to segments of RK2 carrying these two loci, mini-RK2 replicons generated by partial HaeII digestion included a third region separated from $oriV_{RK2}$ by the segment containing Tn1 (18). This region carries a locus, trfB, which is able to complement in trans the temperature-sensitive replication defect of a mutant mini-RK2 derivative (16) and a locus korD, probably identical to trfB, which is able to suppress in trans the plasmid-

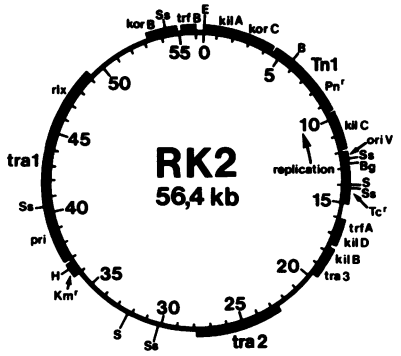

FIG. 1. Map of RK2. Key restriction sites are indicated by: E, EcoRI; B, BamHI; Bg, BglII; H, HindIII; S, SalI; Ss, SstII. Coordinates are in kilobase pairs (kb) from the EcoRI site assuming a total size of 56.4 kb (8); recent results suggest that 60 kb is a better estimate (9). Loci not mentioned in the text are: pri, encoding a DNA primase (10) rlx, the site of the relaxation complex and origin of conjugal transfer (11); tra3 a locus involved in transfer (12); tra1 and tra2, blocks of several loci involved in transfer (13). The loci korA and korD map at the same position as trfB (14).

lethal effect of kilD (14). The kilD locus is closely linked to, but distinct from, trfA and its presence on one of the two HaeII fragments which carry trfA probably dictated the presence of the trfB region in mini-plasmids obtained by HaeII deletion (14, 18). The subsequent deletion of the trfB region from a plasmid carrying the trfA region to form pCT87 (16) probably required the unwitting selection of a spontaneous mutation in the trfA region which effectively inactivated kilD. Other such pairs of kil and kor loci are present on the genome of RK2 (19) and are indicated in Fig. 1. In particular, kilA is suppressed by korA which maps together with trfB and korD and may be the same locus, and kilB, which maps close to trfA and kilD, is suppressed by korB which lies close to trfB/korA/korD (14, 19). The trfB region also contains incC, a locus which expresses incompatibility towards parental type RK2 replicons (K. Ellis & P. Barth, personal communication; 20; 32). We have recently shown both that trfB/korA/korD, incC and korB are transcribed from a common promoter, although korB may have a secondary promoter (32) and that trfA and the cistron encoding a short polypeptide (P_{116}) of unknown function are transcribed from a common promoter (21,22). In this paper we present the nucleotide sequences of these promoters and also that of a region which we believe may be a promoter responsible for the expression of the kilB locus. These sequences contain

potential regulatory sites and, together with experiments on the regulation of transcription in the trfA region reported elsewhere (33) allow us to propose a scheme for the integrated control of a number of loci on RK2.

MATERIALS AND METHODS

Enzymes

Restriction endonucleases, DNA modifying enzymes and RNA polymerase holoenzyme were purchased from Bethesda Research Laboratories, Cambridge Biotechnology Laboratories, Boehringer Mannheim or Amersham International and were used under conditions similar to those recommended by the suppliers.

Sequence Determination

DNA sequences were determined by the method of Maxam and Gilbert (23) with minor modifications. The hot alkali A>C reaction was used in addition to the C-, (C+T)-, (A+G)- and G-specific reactions to improve the certainty of A/G assignments and the incubations for this reaction and for the piperidine cleavage step were carried out with the samples sealed in glass capillaries. Plasmid DNA for sequence determination was prepared as described (22) and digested with suitable restriction endonucleases. The fragments generated were either labelled at the 5′ end by dephosphorylation using bacterial alkaline phosphatase or calf intestinal phosphatase following by rephosphorylation using T4 polynucleotide kinase and $[\gamma^{32}P]$-ATP or labelled at the 3′end using terminal deoxynucleotidyl transferase and $[\alpha^{32}P]$-cordycepin triphosphate or $[\alpha^{32}P]$-dideoxy-ATP. All three nucleotides were from Amersham International at 3,000-5,000 Ci/mmol. The labelled fragments were redigested with different restriction endonucleases to generate singly labelled fragments, which were resolved by agarose gel electrophoresis as described previously (22). The required fragments were recovered by electroelution. All sequences reported were determined on both strands.

Sequencing Strategies

The sequence of the trfB promoter region was determined by 3′ labelling at the HincII and BglI sites (Fig. 2) using pCT16 (18). Attempted 5′ labelling at the HincII site was not successful and so we constructed pCAS160 and pCAS161, consisting respectively of the appropriate HincII to HindIII or EcoRI to HincII fragments of pCT16 inserted into the polylinker of pUC9 (25). We were then able to use 5′ labelling at the EcoRI or HindIII site adjacent to the HincII site in each plasmid to obtain sequence

information for the second strand (Fig. 3).

The sequence of the wild type trfA promoter region was determined by labelling each strand at the EcoRI site of pCT294.6 (22), in which the ClaI site of pBR322, 25 bp (base pairs) from the EcoRI site is fused to a TaqI site within the trfA region and which retains complete copy of the trfB region. The part of the trfA region present in pCT294.6 is indicated in Fig. 3.

The sequences of the mutant trfA promoter region of pCT87 and of the major leftward promoter region of the trfA region were determined by labelling each strand at EcoRI sites with the trfA region in pVI108.1 and in pVI101.1 and pVI106.1 respectively (Fig. 3). These were constructed by mutagenesis of pCT87 with Tn1723, which has EcoRI sites 15 bp from the outer ends of its terminal inverted repeats, followed by deletion of all but 30 bp of the transposon by partial EcoRI digestion and religation (21).

RESULTS AND DISCUSSION

Identification of the trfB Promoter

Fig. 2 shows a map of the trfB/korB region of RK2. We have previously shown by deletion analysis that the promoter responsible for the expression of trfB/korA/korD and at least partly for the expression of korB overlaps the HincII site in the trfB region (14; 32). We confirmed the presence of an RNA polymerase binding site at this position by demonstrating complete, specific, protection of this HincII site from cleavage following preincubation of pCT16 with the polymerase, using a method adapted from reference 26 (data not shown). The strength of this protection was consistent with the results of Wohlleben and Burkardt (27) who mapped RNA polymerase binding sites on RK2 by electron microscopy and found that one of the strongest lay at a position equivalent to that of this HincII site. Fig. 5a shows the nucleotide sequence which we have determined for the region surrounding this HincII site, which shows strong homology to the E.coli promoter concensus sequence (28). The sequence which we report differs at one point from the sequence including most of this region which was recently reported by Bechhofer and Figurski (24); in their sequence one of the three bases numbered 17 to 19 in Fig. 5a is absent. This difference is important because it lies within one of the palindromic sequences discussed below. We are confident that our sequence is correct at this point as we have clear sequencing gels showing each strand. The sequence information to the left of the HincII site reported by Bechhofer and

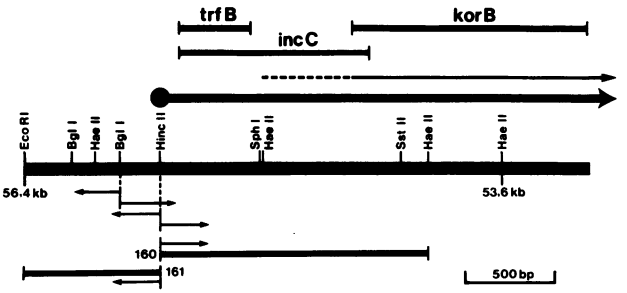

FIG. 2. Map of the trfB/korB region. Below the restriction map are indicated the segments of the region present in pCAS160 and pCAS161, and the directions and approximate extent of sequence determination from each site. Above the map are shown the direction and extent of transcription from the trfB promoter and from a probable secondary promoter within the region, and the approximate coding regions associated with the trfB, incC and korB loci (14, 24, 32).

Figurski was derived from only one strand (24) and, apparently, by labelling at a site some 310 bases from this point. It is therefore more likely that the run of three identical bases appeared to be only two bases on their sequencing gels than that there is a mutational difference between the plasmids from which the sequence was derived.

Promoters in the trfA Region

Fig. 3 shows a map of the trfA region of RK2 on which are indicated the positions of known coding sequences (22), the position of certain Tn1723 insertions (21) and the direction and relative strength of transcription in

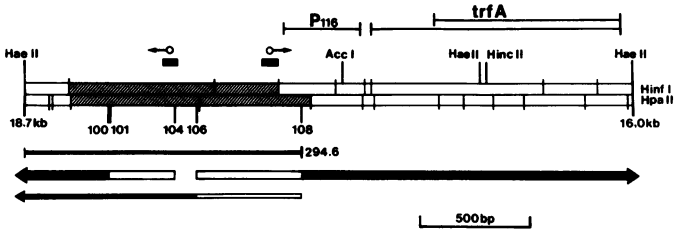

FIG. 3. Map of the trfA region. Above the restriction map are indicated known coding regions (21,22) and the positions of the promoters identified in this work. Hatched regions of the map indicate HinfI and HpaII fragments whih are strongly bound by RNA polymerase. Below the map are indicated the points of insertion of Tn1723 in various derivatives of pCT87 (21), the segment of the region present in pCT294.6 (14,22) and the directions and possible extents of transcriptions identified in the trfA region (21; 33). The regions within which these transcriptions must originate are indicated by the unfilled sections.

various regions (21; 33). Since the insertion of Tn1723 in pVI108 has a polar effect on the expression of trfA while the insertion in pVI106 has no effect (21) the promoter for trfA must lie between the sites of these two insertions. There are two discrete levels of leftward transcription in the trfA region, a low level originating between the same insertions as the rightward (trfA) transcription and a higher level originating between the insertion points in pVI100 and pVI104. We have recently shown (33) that the level of the rightward transcription is reduced when trfB (korA/korD) is present in trans and is further reduced when both trfB and korB are present, while the levels of the leftward transcriptions are not affected by trfB alone but are reduced when both trfB and korB are present. We have preliminary results which suggest that korB is at least partly able to control these transcriptions in the absence of trfB. The regulatory properties of the leftward transcriptions suggested that they might be responsible for the expression of the kilB locus. Although we have shown that this locus is not completely present within the two HaeII fragments which carry trfA (14) as was originally proposed (19), it is possible that part of it is present, as a maximum of only 900 bp of additonal RK2 DNA adjacent to the trfA region are required for the expression of kilB (14).

In order to obtain evidence for the positions of RNA polymerase binding sites within the trfA region we end-labelled HinfI and HpaII digests of the trfA region and determined which fragments were retained on nitrocellulose filters in the presence of RNA polymerase, essentially according to the procedure described in reference 29. Two fragments from each digest were strongly bound and the positions of these (Fig. 3) indicate the presence of two RNA polymerase binding sites within regions which correspond to those in which the rightward (trfA) and the stronger of the leftward transcriptions originate.

Identification of the trfA Promoter Sequence

In order to identify more precisely the position of the trfA promoter we made use of the DNAase footprinting technique (30). Samples of a singly end-labelled fragment containing the trfA promoter region were partially digested with DNAase I in the presence or absence of RNA polymerase and run on sequencing gels beside samples subjected to the normal sequencing reactions. We observed a single region of clear protection covering the region from 25 to 70 bp upstream from the start of the 116 codon open reading frame which preceeds that of trfA (Figs. 3 & 4). The protected

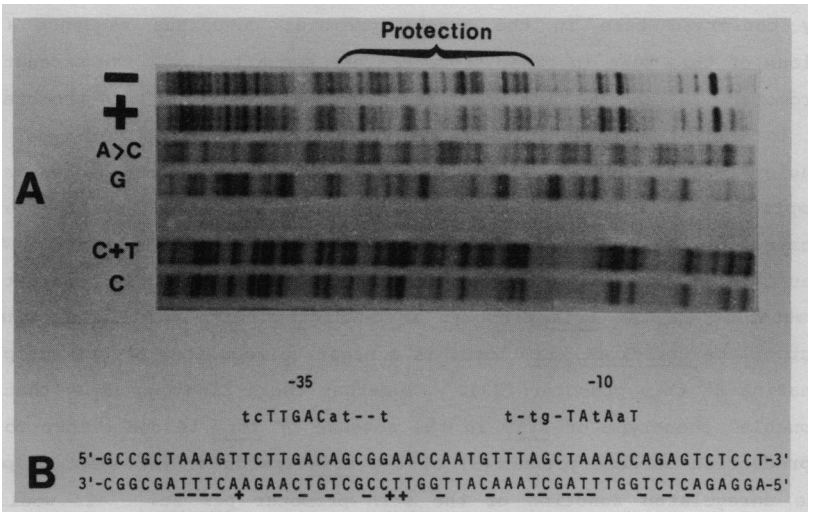

FIG. 4. RNA polymerase binding in the trfA promoter region. A; 8% sequencing gel. pVI108.1 DNA was 5′ labelled at the EcoRI site left by deletion of the Tn1723 insertion from pVI108 (Fig. 5). +; partial DNAse I digestion in the presence of a 50-fold molar excess of RNA polymerase (30). -; similar digestion in the absence of polymerase. B; Sequence of the trfA promoter region of pCT87, indicating sites of reduced (-) or enhanced (+) cleavage in the presence of RNA polymerase and regions of homology to the E.coli concensus sequence (28).

region contains considerable homology to the E.coli promoter concensus sequence (28).

The trfA Promoter of pCT87(kilD⁻) is Mutant

We have determined the nucleotide sequence of the entire trfA region (22; C.A. Smith, unpublished). All but 12 bp of the sequence was determined on at least one strand using derivatives of RK2 which had always retained a copy of the trfB region, as it appeared that derivatives lacking this region must carry a mutation within the trfA region which effectively inactivates the kilD locus (14, 22). In addition the sequence of most of the minimal segment known to express a KilD phenotype in the wild-type was determined from derivatives of pCT87, the first mini-RK2 derivative constructed carrying the trfA region but lacking essentially all of the trfB region (16). The only sequence difference which we have observed between the wild-type trfA region and that of pCT87 lies in the putative Pribnow box of the trfA promoter sequence which reads 5′-TAAACT in the wild-type (Fig. 5b) but 5′-TAAACC in pCT87 (Fig. 4). This mutation not only decreases the homology of the Pribnow box to the concensus 5′-TAtAaT but alters the most

highly conserved base in the E.coli promoter concensus sequence (28). Mutations of this base in other promoters are known to lower the strength of the promoter (28) and we have preliminary evidence that the trfA promoter of pCT87 is indeed considerably weaker than that of the wild-type (33; V.Shingler, A.A.K. Hussain & C.M. Thomas, unpublished), although comparison is complicated by our inability thusfar to clone the wild-type promoter in the absence of the trfB/korA/korD locus. The finding that a mutation which enables the trfA region to be cloned in the absence of korD is apparently a down mutation in the trfA promoter is consistent with our finding that a product of the trfB/korA/korD locus is a negative regulator of transcription originating at this promoter (33). Together these findings imply that the "unclonable" phenotype of kilD in the absence of korD is due either to the over-production of some product(s) of the trfA operon or (at least in part) to the unregulated function of the trfA promoter per se. It would be tempting to suggest that the first product of the trfA operon, P_{116}, which is synthesized at high level and has no known function (21,22) is associated with kilD. However, the trfA region of one of our deletion derivatives, pCT294.6 (Fig. 3), which shows this phenotype retains only the first 30 codons of the P_{116} cistron (14, 22). It is therefore possible that, at least in some cases, the KilD phenotype may be due to the direct effect of a strong unregulated transcription on the replication of the vector plasmid. Such an inhibitory effect has been observed for pMB1 replicons (31).

The sequence of the wild-type trfA promoter region shown in Fig. 5b overlaps by 10 bases with the sequence including the trfA gene which we have reported previously (22).

Identification of a Promoter Regulated by korB

Examination of the nucleotide sequence of the second segment of the trfA region which contains an RNA polymerase binding site revealed a single promoter-like sequence (Fig. 5c), of fair homology to the E.coli promoter concensus sequence, at the position indicated in Fig. 3. The position of this sequence is consistent with its representing the promoter responsible for the stronger level of leftward transcription (Fig. 3). As the position and regulatory properties of this transcription are consistent with its being responsible for the expression of the kilB locus, we tentatively identify this promoter-like sequence as the major promoter for kilB. The segment of the trfA region within which the weaker leftward transcription originates (Fig. 3) contains more than one sequence of weak homology to the E.coli promoter concensus sequence (data not shown) and further mapping will

be required to identify the promoter(s) responsible for this weak transcription.

Comparison of the trfB, trfA and Putative kilB Promoter Regions

Fig. 5 shows the sequences of the trfB, trfA and putative kilB promoter regions aligned to show their homology to the E.coli promoter concensus sequence (28). That of the trfB promoter shows the greatest homology, with 5/6 bases homologous to the -35 box and 4/6 homologous to the -10 Pribnow box, the non-concensus bases in each case being at the most weakly conserved positions (28). The spacing of 17 bases between the -35 and -10 boxes is optimal (28) as noted by Bechnofer and Figurski (24), and there is a perfect CAT box 5 bases beyond the -10 box at a suitable site for mRNA initiation. This high degree of homology suggests a very powerful promoter, and we have preliminary evidence that this is the case (C.A. Smith & C.M. Thomas, unpublished).

The trfA promoter region contains an ideal concensus -35 box and in the wild type has the same -10 box with 4/6 homology to the concensus sequence as the trfB promoter. However, in this case the spacing is 18 bases and there is no purine ideally positioned to serve as the mRNA initiation site. It is therefore likely that this promoter is somewhat weaker than the trfB promoter.

The putative kilB promoter sequence shows poorer homology to the concensus sequence, with the concensus base at only the three most highly conserved positions in the -35 box (28) and a -10 box with 4/6 homology to the concensus, but with C in place of the final T which is the most highly conserved base of the E.coli concensus sequnce (28). The spacing between the -35 and -10 boxes is 16 bases and there is again no suitably situated purine to serve as the initiation point for mRNA synthesis. This suggests that this promoter is probably weaker than the wild-type trfA promoter, although it is known to be slightly stronger than the mutant trfA promoter of pCT87 (33).

The Three Promoter Regions Share Possible Operator Sites

It is apparent from Fig. 5 that the trfB and trfA promoter regions show much more sequence homology than is accounted for by their mutual homology to the promoter concensus sequence. Most of this additional homology is accounted for by the presence in each of two palindromic sequences. One of these (5′-GTTTAGCTAAAC) overlaps with the -10 Pribnow boxes, while the second (5′-TTTAGC$^{G}_{C}$GCTAAA) lies just upstream from the -35 boxes. The positions of these palindromes suggest that either or both

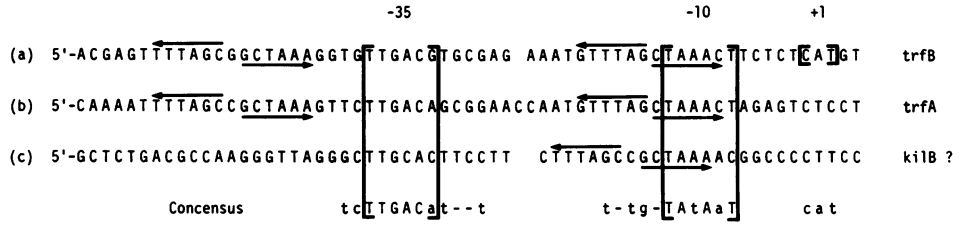

FIG. 5. Comparison of promoter regions The sequences of the non-transcribed strands of the three promoter regions identified are aligned to show homology to each other and to the E.coli concensus promoter sequence (28). The wild-type trfA promoter region is shown. The -35, -10 and (when present) +1 boxes are marked and the palindromic regions are indicated. The less highly conserved bases in the concensus sequence are indicated by lower case letters.

could serve as operator site(s) for negative control of these promoters by a repressor or repressors, while the second might alternatively (or in addition) serve as a site for positive regulation. It is noteworthy that the two palindromic sequences are closely related, each including invert repetition of the sequence 5'-TTTAG, although with a different spacing. Furthermore, the putative kilB promoter sequence also contains a copy of the second palindromic sequence, but in this case it overlaps with the Pribnow box (Fig. 5c), thus occupying a position equivalent to that of the first palindromic sequence in the trfA and trfB promoters.

Thus two operator-like sequences are present in the trfA promoter region, transcription from which is controlled negatively by both trfB (korA/korD) and korB, one of which is also present in the putative kilB promoter region, transcription from which is not controlled by trfB alone but only by korB (possibly in combination with trfB). We therefore postulate that the palindromic sequence which is present only in the trfA and trfB promoters is a binding site for a product of the trfB/korA/korD locus, while the palindrome present in all three promoter regions may be a binding site for the korB product, either alone or in combination with the trfB product. On this basis we would predict that the trfB promoter is (auto)-regulated by products both of trfB/korA/korD and of korB. We have some preliminary evidence that this may be the case (C.A. Smith & C.M. Thomas, unpublished). If, as seems probable, the same product of the trfB/korA/korD region is responsible for the suppression both of kilD and of kilA, we would further predict that the kilA locus is transcribed from a

promoter region which contains a palindromic sequence homolous to 5'-GTTTAGCTAAAC.

Knowledge of the sequences of these three promoter regions, together with the results of our studies on the regulation of transcription in the trfA region (33), has considerably advanced our knowledge of the control of the trfA, kilB and kilD loci of RK2 and enabled us to make testable predictions concerning the control of the trfB, korB and kilA loci. The involvement of both trfB/korA/korD and korB in the regulation of trfA, a locus essential to the replication system of RK2 (16) may account for the presence of loci functionally interchangeable with, and presumably homologous to, both korA and korB on diverse and apparently distantly related IncP plasmids (19). The apparently coordinated control of trfA and at least some of the kil loci of RK2 may imply that the latter are also involved in some (non-essential) aspect of plasmid maintenance (e.g. coordination of cell division with plasmid replication) or that high expression of both trfA and the kil loci is required under certain circumstances (e.g. following conjugal transfer). These possibilities are not mutually exclusive and much work remains to be done on the control circuits of this broad host range plasmid.

ACKNOWLEDGEMENTS

We thank Mrs. E. Badger for typing the manuscript and Mr. B. Price for photographing the figures. This work was partly financed and C.A. Smith was supported by M.R.C. project grant no. G80/0374/9CB awarded to C.M. Thomas. V. Shingler was supported by an M.R.C. Research Studentship.

*To whom correspondence should be sent

REFERENCES

1. Datta, N. and Hedges, R.W. (1972) J. Gen. Microbiol. 70, 453-460.
2. Olsen, R.H. and Shipley, P. (1973) J. Bacteriol. 113, 772-780.
3. Beringer, J.E. (1974) J. Gen. Microbiol. 84, 188-198.
4. Cho, J.J., Panopoulos, N.J. and Schroth, M.N. (1975) J. Bacteriol. 122, 192-198.
5. Windass, J.D., Worsey, M.J., Pioli, E.M., Pioli, D. Barth, P.T., Atherton, K.T., Dart, E.C., Byrom, D., Powell, K. and Senior, P.J. (1980) Nature 287, 396-401.
6. Burkardt, H.-J., Reiss, G. and Pühler, A. (1979) J. Gen. Microbiol. 114, 341-348.
7. Stokes, H.W., Moore, R.J. and Krishnapillai, V. (1981) Plasmid 8, 202-212.
8. Meyer, R., Figurski, D. and Helinksi, D.R. (1977) Mol. Gen. Genet. 152, 129-135.
9. Lanka, E., Lurz, R. and Fürste, J.P. (1983) Plasmid 10, 303-307.
10. Lanka, E. and Barth, P.T. (1981) J. Bacteriol. 148, 769-781.

11. Guiney, D.G. and Helinski, D.R. (1979) Mol. Gen. Genet. 176, 183-189.
12. Barth, P.T. (1979) in Plasmids of medical, environmental and commercial importance, Timmis, K.N. and Pühler, A., Eds., pp 399-410. Elsevier/North Holland, Amsterdam.
13. Barth, P.T., Grinter, N.J. and Bradley, D.E. (1978) J. Bacteriol 133, 43-52.
14. Smith, C.A. and Thomas, C.M. (1983) Mol. Gen. Genet. 190, 245-254.
15. Meyer, R. and Helinski, D.R. (1977) Biochim. Biophys. Acta 478, 109-113.
16. Thomas, C.M. (1981) Plasmid 5, 277-291.
17. Schmidhauser, T.J., Filutowicz, M. and Helinski, D.R. (1983) Plasmid 9, 325-330.
18. Thomas, C.M., Meyer, R. and Helinski, D.R. (1980) J. Bacteriol. 141, 213-222.
19. Figurski, D.H., Pohlman, R.F., Bechhofer, D.H., Prince, A.S. and Kelton, C.A. (1982) Proc. Nat. Acad. Sci. USA 79, 1935-1939.
20. Meyer, R. and Hinds, M. (1982) J. Bacteriol. 152, 1078-1090.
21. Shingler, V. and Thomas, C.M. (1984) J. Mol. Biol. In press.
22. Smith, C.A. and Thomas, C.M. (1984) J. Mol. Biol. In press.
23. Maxam, A.M. and Gilbert, W. (1980) in Methods in Enzymology Vol 65, Grossman, L. and Moldave, K. Eds., pp 499-560. Academic Press, New York.
24. Bechhofer, D.H. and Figurski, D.H. (1983) Nucleic Acids Res. 11, 7453-7469.
25. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
26. Gunsalus, R.P. and Yanofsky, C. (1980) Proc. Nat. Acad. Sci. USA 77, 7117-7121.
27. Burkardt, H.-J. and Wohlleben, W. (1981) J. Gen. Microbiol. 125, 189-193.
28. Hawley, D.K. and McClure, W. (1983) Nucleic Acids Res. 11, 2237-2255.
29. Seeburg, P.H. and Schaller, H. (1975) J. Mol. Biol. 92, 261-277.
30. Galas, D.J. and Schmitz, A. (1978) Nucleic Acids Res. 9, 3157-3170.
31. Stueber, D. and Bujard, H. (1982) EMBO J. 1, 1399-1404.
32. Smith, C.A. and Thomas, C.M. (1984) J. Gen. Microbiol. In press.
33. Shingler, V. and Thomas, C.M. (1984) Mol. Gen. Genet. In press.