

The *c4* repressor of bacteriophage P1 is a processed 77 base antisense RNA

Martin Citron and Heinz Schuster*

Max-Planck Institut für Molekulare Genetik, Ihnestr. 73, D-1000 Berlin 33, Germany

Received March 6, 1992; Revised and Accepted May 18, 1992

GenBank accession no. M35139

ABSTRACT

The *c4* repressors of the temperate bacteriophages P1 and P7 inhibit antirepressor synthesis and are essential for establishment and maintenance of lysogeny. Using *in vivo* complementation tests we have previously shown that *c4* is an antisense RNA acting on a target, *ant* mRNA, which is transcribed from the same promoter. Here we identify the *c4* repressor molecule of P1 as a 77 ± 1 base RNA by mapping its termini and show that the *c4* RNA in P7 lysogens has the same or a similar size. P1 *c4* RNA is encoded in a region shown to be sufficient for *c4* complementation. It covers exactly the 74 bases previously suggested to fold into a stem-loop secondary structure essential for *c4* function. Furthermore, we demonstrate that the 5' end of *c4* RNA is generated by processing. Thus, *c4* is the first example of an antisense RNA to be processed. A possible mechanism of processing is discussed.

INTRODUCTION

During recent years, antisense inhibition of gene expression has been demonstrated in a variety of systems, and several antisense RNAs in prokaryotes have been analyzed in detail (for reviews see 1, 2). We have recently described a novel antisense system in the immunity region *immI* of the temperate bacteriophages P1 and P7 (for all but selected references on P1 and P7 see 3). Here, the *c4* repressor is an antisense RNA which acts on the target antirepressor (*ant*) mRNA. By this means antirepressor synthesis is repressed in P1 and P7 lysogens. The novelty is that *c4* and *ant* mRNA are transcribed from the same promoter (4).

The P1 *immI* elements have been described previously (5, 6). *ant* encodes an antirepressor, and two antirepressor proteins *ant1* and *ant2* are translated from a single open reading frame. An open reading frame *orfx* of unknown function overlaps the start of *ant1* (Fig. 1). Its function is required for *ant* expression. The *c4* gene is located upstream of *orfx* and encodes the *c4* repressor which prevents expression of *ant*. A tandem promoter P51a/b and a *c1* repressor-controlled operator, Op51, overlapping the -35 region of P51a, are located in front of the *c4* gene (7, 5). It is important to note that *c4*, *orfx*, and *ant1/2* are cotranscribed in that order, and that transcription starting from P51b is sufficient to express *c4* and *ant*, the latter only if *c4* cannot act.

In *in vivo* complementation tests, *c4* activity of P1 could be reduced to a 85 bp fragment, the 'c4 mini' region (Fig. 1). This region does not encode a protein but must be transcribed to obtain *c4* complementation. Similarly, *c4* activity of P7 could be reduced to a 139 bp fragment. Therefore, the *c4* repressors of bacteriophages P1 and P7 are not proteins but regulatory RNAs. Each RNA contains two short sequence elements a' and b' that are complementary to their target sequences a2 and b2 in *ant* mRNA (Fig. 1). We have shown that this complementarity is essential for *c4* action and proposed that *c4* acts as a translational repressor of *orfx*, thereby also inhibiting *ant* expression (4). P1 and P7 are heteroimmune (8, 9) and the immunity difference was mapped to the *c4* region (8). It follows from our results that the heteroimmunity is due to just two substitutions in each of the complementary sequences of *c4* and *ant* mRNA (Fig. 1). Since *c4* acts in *trans* on a target RNA transcribed from the same promoter, we have classified *c4* as a new type of antisense RNA. We have also suggested a secondary structure model for the transcript of the *c4* mini region with the complementary regions in loops as important sites for antisense control (4).

However, the existence of *c4* RNA has not yet been proven directly. In this paper we demonstrate the presence of a *c4* RNA in P1- and P7-lysogenic bacteria. By mapping its termini we show that the *c4* RNA of P1 is a 77 ± 1 b long transcript which overlaps the *c4* mini region. We also find a *c4* RNA of similar or the same size in P7 lysogens. Thus, *c4* RNA is the first example of an antisense RNA which originates from a longer transcript by a processing event.

MATERIALS AND METHODS

Bacterial strains and bacteriophages

The *E. coli* K12 strains (with relevant markers) used were C600 *supE44* (10), SCS1 *supE44 recA1* (Stratagene, Inc.), the P1Cm-, P7- and P7c4.2ts lysogens of C600, and the P1Cm^{vir}*ant16*-, P1Cm^{vir}*ant10*- and P1Cm^{c4.32ant17} lysogens of YMC (11). Bacteriophages used were P1Cm (12), P7 (13) and P7c4.2ts (14).

Plasmid constructions

The cleavage sites of restriction enzymes and deletion endpoints used for creating P1 DNA fragments are shown in Fig. 1. Numbers indicate the P1 nucleotide positions shown in Fig. 1.

* To whom correspondence should be addressed

generated by *in vitro* transcription of linearized plasmids by T7 RNA polymerase. Samples were run on a 6% polyacrylamide sequencing gel in 8 M urea.

RESULTS

c4 RNA is detected in P1- and P7-lysogenic bacteria

In order to demonstrate *c4* RNA *in vivo* we performed a Northern blot analysis of several P1 lysogens. We used a *c4* probe which covers the entire region from the promoter P51a to the start of the *orfx* gene (Fig. 1, position 4–410). As expected, no cross-hybridization was detected with RNA from nonlysogenic bacteria used as control (Fig. 2A, lane 1). In contrast, a strong and a rather weak band, about 75 b and 200 b in size, respectively, are observed in lysogens of P1 (Fig. 2A, lane 3), P1*vir^sant16* (lane 4) and P1*vir^sant10* (lane 5). These bands are not detectable in a P1*ant17c4.32* lysogen (Fig. 2A, lane 2). We consider the 75 b and 200 b RNAs to be *c4* specific since they appear only in lysogens that produce active *c4* repressor, as shown by their immunity to P1 superinfection. The *vir^sant* double mutants express active *c4* RNA because the *vir^s* mutation (leading to constitutive expression of *orfx-ant*) as well as the *vir^s* suppressor mutations *ant16* and *ant10* (which abolish *orfx-ant* synthesis) are located downstream of an unchanged *c4* gene (Fig. 1). In contrast, in a P1*ant17c4.32* lysogen *c4* is defective and transcription of the *immI* region is reduced drastically due to the P51 promoter down mutation *ant17*. The failure to detect a transcript of a P1*ant17c4.32* lysogen also demonstrates the absence of other promoters within the first 410 bp of the opposite DNA strand.

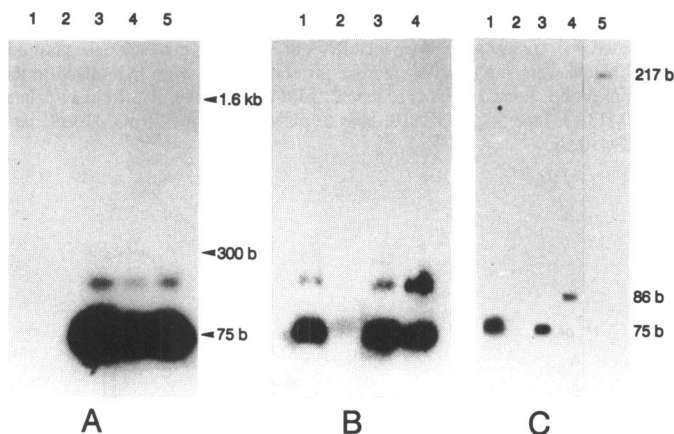


Figure 2. Northern blot analysis of *c4* RNA from P1- and P7 lysogens. Total RNA from P1- and P7 lysogens and C600 as a control was subjected to gel electrophoresis, and the RNAs were hybridized to different probes. (A) 21 µg RNA from C600 (lane 1), and from the lysogens of P1 *ant17c4.32* (lane 2), P1 (lane 3), P1*vir^sant16* (lane 4), and P1*vir^sant10* (lane 5). Samples were run in a 2% agarose glyoxal gel and were hybridized to a probe covering positions 4–410 of P1 *immI* (Fig. 1). Approximate fragment lengths indicated at the right were determined from migration of 16S rRNA, a ΦX174 *Hae*III fragment, and tRNA in descending order. (B) 40 µg of total RNA from C600 (lane 2) and the lysogens of P1 (lane 1), P7 (lane 3) and P7*c4.2ts* (lane 4) grown at 30°C were run in a 2% agarose glyoxal gel and were hybridized to a probe covering positions 20–410 of P7 *immI* (see Fig. 1). (C) 20 µg RNA from C600 P1 (lane 1) and C600 (lane 2) were separated in a 6% polyacrylamide gel containing 8 M urea before hybridization to the *c4* mini probe covering positions 242–326 of P1 *immI* (see Fig. 1). Labeled marker transcripts of the length indicated at the right were run in parallel (lane 3–5).

The closely related but heteroimmune bacteriophage P7 expresses a *c4* repressor of different sequence specificity which results from just two substitutions in each of the complementary sequences of *c4* and *ant* mRNA (Fig. 1). Therefore, we expected to detect an RNA of similar or the same size as in P1 also in P7 lysogens. Indeed, using a P7 *c4* probe covering the entire region from P51a to the start of *orfx* (20–410) we could show a band pattern for P7 (Fig. 2B, lane 3) and P7*c4.2ts* (lane 4) which is similar to that of P1 (lane 1). Only the amount of the 200 b band is increased in P7*c4.2ts* relative to P7. We conclude that both P1 and P7 express a major *c4* RNA of about 75 b in size. To determine the size of P1 *c4* RNA more precisely, RNA from C600(P1) (Fig. 2C, lane 1) and C600 (lane 2) was separated on a denaturing polyacrylamide gel alongside labeled marker transcripts of known length (lanes 3 to 5). Hybridization was then performed with a '*c4* mini' probe (242–326) which covers only the '*c4* mini' region previously found to be sufficient for *c4* complementation (4). The result indicates that P1 *c4* RNA must be slightly longer than 75 b and considerably shorter than 86 b.

The *c4* repressor of P1 is a 77 ± 1 base antisense RNA

In order to determine the 5' end of *c4* RNA exactly, we performed a primer extension analysis. Total RNA from a P1

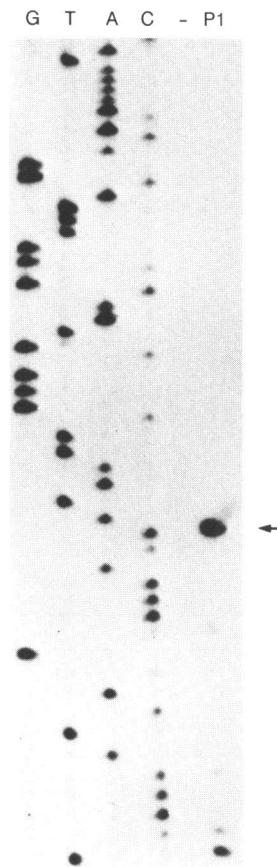


Figure 3. Mapping of the *c4* RNA 5'-terminus by primer extension. Total RNA from C600 (P1) and C600 (–) as control was hybridized to the oligonucleotide primer (see Fig. 1), extended with reverse transcriptase, and run in a 6% polyacrylamide sequencing gel. The same gel contained DNA sequencing reactions (G, T, A, and C, respectively) using the same primer and plasmid pUM10 (4) containing a P1 *Dra*I–*Eco*RI restriction fragment (Fig. 1, position 4–410). The arrow marks the base C which is complementary to base G at the 5'-end of *c4* RNA.

lysogen was hybridized to an oligonucleotide primer (Fig. 1) and extended with reverse transcriptase. The cDNA product was then compared with the DNA sequence ladder obtained from the sequencing reactions with the same primer and plasmid pUM10 containing the entire P1 *c4* region (4–410; Fig. 1). The 5' end of *c4* RNA was found to be the guanine at position 253 (Fig. 3). This finding is in accordance with results of *c4* complementation studies *in vivo* which indicated that transcription of a functional *c4* RNA can start anywhere upstream of the $\Delta 4$ deletion endpoint (4) (position 254, Fig. 1).

RNase protection experiments were done in order to determine the exact size and the 3' terminus of *c4* RNA. Total RNA from a P1 lysogen was hybridized to complementary transcripts as shown in the inset of Fig. 4. The hybrids were digested with RNase A and RNase T1 and electrophoresed (Fig. 4). With one exception, only the transcripts which covered the *c4* mini region totally produced a hybrid of 77 ± 1 bp in length. The exception is probe $\Delta 6$ (4–326) which yielded a product about 3 bp shorter (Fig. 4, lane 14). This indicates that the 3' end of the *c4* RNA is 3 b longer than the *c4* mini region determined genetically (Fig. 1). The RNase protection results are in accordance with the 5' end of *c4* RNA as determined by primer extension and the size of *c4* RNA as determined by Northern blotting. Taken

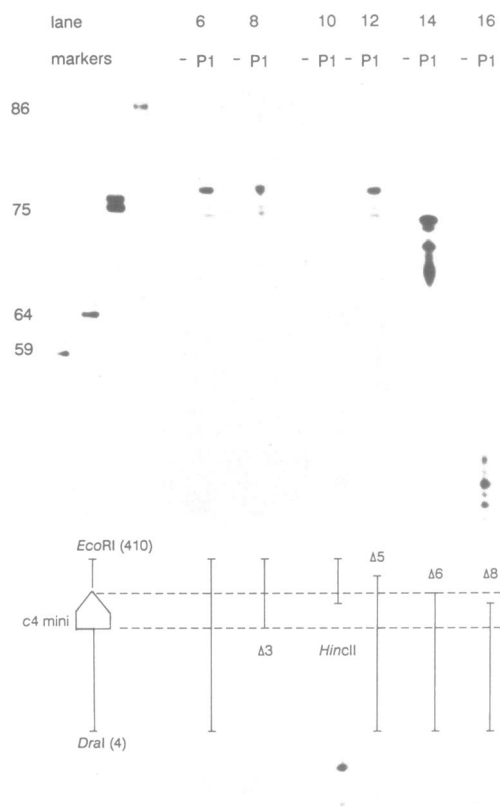


Figure 4. Mapping of *c4* RNA by ribonuclease protection analysis. Total RNA from C600 (P1) and C600 (–) as a control was hybridized to different RNA probes, digested with RNase A and RNase T1 for 60 min, and separated by electrophoresis on a 6% polyacrylamide sequencing gel. Labeled marker transcripts of known length were run in parallel. In the inset, the size of the RNA probes is shown by a vertical line in relation to the *c4-orfX* regulatory region. The probes are complementary to the strand shown in Fig. 1 in which restriction enzyme cutting sites and deletion endpoints are indicated.

together, we conclude that *c4* RNA is coded by positions 253 through 329 of *immI*.

c4 RNA is a processed transcript

The results described above indicate that *c4* RNA is a processed transcript, since it starts 113 bases downstream from the promoter

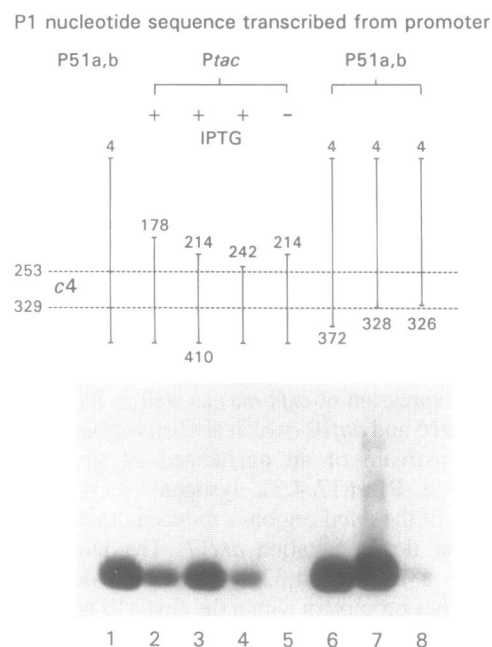


Figure 5. Northern blot analysis of *c4* RNA from plasmids. (Upper part) The size of the P1 DNA insert (with nucleotide numbers as in Fig. 1) is shown by a vertical line. The promoter from which the insert is transcribed is indicated above the line. (Lower part) 30 μ g total RNA of SCS1 cells containing the plasmids listed below were run in a 2% agarose glyoxal gel and were hybridized to the *c4* mini probe. Lane 1: pUM11, lane 2: pJM13 Δ 11, lane 3: pJM13 Δ 10, lane 4: pJM13 Δ 3, lane 5: pJM13 Δ 10, lane 6: pUM10 Δ 5, lane 7: pUM10 Δ 9, lane 8: pUM10 Δ 6.

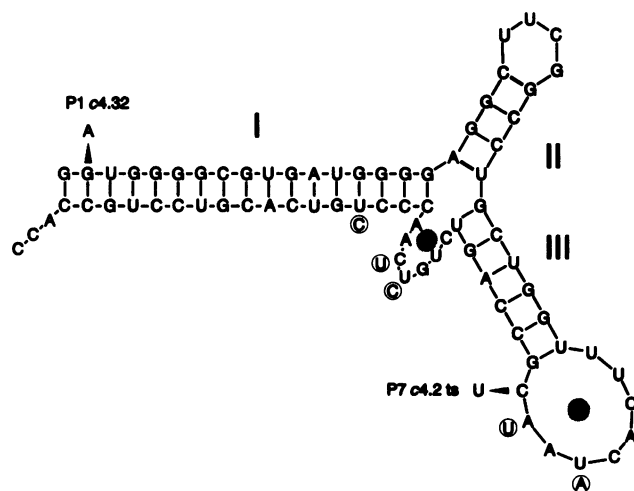


Figure 6. Secondary structure model of the *c4* antisense RNA. The sequence of *c4* RNA as determined in this paper (position 253 to 329, Fig. 1) was folded as described previously (4). Positions where the sequence of P7 differs from that of P1 are marked by circled letters. Mutations are indicated by arrowheads, followed by the name of the phage. Stems are marked by Roman numerals.

P51b, deletion of which abolishes *c4* complementation *in vivo* (7, 4). We wanted to substantiate these findings and to investigate whether P1 specific sequences upstream and downstream of *c4* are required for processing and termination. For that purpose, we analyzed the transcription of *c4* RNA from several plasmids by Northern blotting. Transcription was initiated from either the P1 promoter P51a,b or the IPTG-inducible *tac* promoter (Fig. 5). Synthesis of *c4* RNA is observed with all plasmids, irrespective of whether the natural P1 sequence upstream of the *c4* mini region is present or not (Fig. 5, lanes 1–4). The *c4* RNA from plasmid pUM11 (Fig. 5, lane 1) has the same size as the *c4* RNA from a P1 lysogen as was found in a parallel run (data not shown). If the *tac* promoter remains repressed (in the absence of IPTG), *c4* RNA is not detectable (Fig. 5, lane 5). Likewise, *c4* RNA is not found when P51a,b is deleted and not substituted by a heterologous promoter (plasmid pUM13 (4), data not shown). These results clearly demonstrate that (i) *c4* RNA is a processed transcript, because there is no P1 specific promoter detectable which could lead to a transcription start at position 253, (ii) sequences upstream of the *c4* mini region are not essential for the 5' end processing of *c4* RNA, nor is there any other bacteriophage specific function or factor required, and (iii) transcription from a P1-specific or heterologous promoter is sufficient for *c4* RNA synthesis which in turn is sufficient for *c4* complementation.

Apparently, the natural P1 sequences downstream of the *c4* mini region are not required for the processing or termination event, because plasmids pUM10Δ5 (Fig.5, lane 6), pUM10Δ9 (lane 7), and pUM10Δ6 (lane 8) synthesize *c4* RNA and they are active in *c4* complementation (4). However, when the 3' terminal nucleotide (position 329) of *c4* RNA is deleted, a smear of *c4* specific RNA of increasing size is always observed (pUM10Δ9, lane 7). Likewise, deletion of three nucleotides from the 3' end of *c4* RNA results in a considerably decreased rate of *c4* RNA synthesis (pUM10Δ6, lane 8). These results indicate that processing or termination is disturbed to some extent if *c4* RNA is not intact.

DISCUSSION

In this study we have analyzed the *c4* RNAs of bacteriophages P1 and P7. The *c4* repressor of P1 was identified as a 77 ± 1 base RNA extending from position 253 to 329 (Fig. 1). An RNA of the same size was detected in P7 by Northern blotting and we assume that its 5' and 3' termini are the same as in P1. Therefore, we conclude that the 77 ± 1 base transcript is the active *c4* repressor molecule. Another RNA, about 200 bases in size, was detected in P1- and P7 lysogens in low amounts. This RNA is, like *c4*, not observed when the P1 prophage carries the P51b promoter-down mutation *ant17* (Fig.2). Therefore, we assume that the 200 base transcript is a *c4* precursor RNA or processing intermediate.

The *c4* RNA overlaps most of the *c4* mini region that was found to be sufficient for *c4* complementation *in vivo*. We have previously suggested that the RNA sequence from positions 253 to 326 folds into a stem-loop secondary structure in which the antisense boxes b' and a' are exposed in loops. This sequence appeared to be absolutely necessary for *c4* function (4). It is striking that *c4* RNA indeed starts at position 253 and terminates just 3 bases beyond the end of the suggested stem at position 329 (Fig. 6). This strongly supports our proposed secondary

structure model. In P1c4.32 a G→A base exchange at the extreme end of the long stem I (Fig. 6) results in a *c4* defective phenotype (7). In preliminary *in vivo* experiments it was found that a *c4.32*-carrying plasmid synthesized much less *c4* RNA than the corresponding *c4*⁺ plasmid (unpublished results). Accordingly, the defectiveness of P1c4.32 might result from an abnormally high turnover rate of the *c4.32* precursor and/or mature RNA. With the start and end of *c4* RNA now being known, it will also be possible to design *in vitro* experiments to analyze the interaction between *c4* and its target RNA. This will help to understand how *c4* can efficiently interact with its target in spite of the short range of complementarity which is restricted to the a'-a2- and b'-b2-regions.

We have shown that *c4* RNA is processed at the 5' end. Whether the 3' end is generated by processing or termination is not known. This adds to the special character of *c4* RNA. It is not only the first example of an antisense RNA to be transcribed from the same promoter as its target RNA, but there is also no other antisense RNA known to be processed at the 5' end (1, 2). This makes *c4* an interesting model for designing artificial antisense RNAs. Sequences upstream of the processing site can be manipulated and do not influence the antisense RNA as long as transcription of the essential region is maintained.

Yet another aspect is noteworthy. In the *c4* RNA model the structure of stem I with G at the 5' end and a single-stranded ACC protruding from the 3' end (Fig. 6) is strikingly similar to the overall structures of the acceptor stem in tRNAs or the hairpin in 4.5 S RNA of *E. coli*. The latter two classes of RNA molecules originate from larger precursors which are processed by RNase P (23, 24). It will be interesting to find out whether the bacteriophages P1 and P7 have exploited the *E. coli* processing system for tRNA and rRNA for their own *c4* RNA.

ACKNOWLEDGEMENTS

H.S. dedicates this manuscript to Werner Schäfer on the occasion of his 80th birthday. We appreciate the expert technical assistance of Susanne Freier. We thank Rolf Bald and Dietmar Vogt for the preparation of oligonucleotides and plasmid DNA, respectively and R.Brimacombe for suggestions to improve writing of the manuscript. We also thank J.Scott and M.Yarmolinsky for providing bacteriophage P7c4.2ts and the P1-lysogenic strains, respectively. This work was supported by Fonds der Chemischen Industrie.

REFERENCES

1. Simons,R.W., and Kleckner,N. (1988) *Annu. Rev. Genet.* **22**, 567–600.
2. Takayama,K.M., and Inouye,M. (1990) *Crit. Rev. Biochem. Mol. Biol.* **25**, 155–184.
3. Yarmolinsky,M.B., and Sternberg,N. (1988) In Calendar,R. (ed.) *The Bacteriophages*. Plenum Press, New York, Vol. 1, pp. 291–438.
4. Citron,M., and Schuster,H. (1990) *Cell* **62**, 591–598.
5. Heisig,A., Riedel,H.D., Dobrinski,B., Lurz,R., and Schuster,H. (1989) *J. Mol. Biol.* **209**, 525–538.
6. Hansen,E.B. (1989) *J. Mol. Biol.* **207**, 135–149.
7. Baumstark,B.R., and Scott,J.R. (1987) *Virology* **156**, 197–203.
8. Chesney,R.H., and Scott,J.R. (1975) *Virology* **67**, 375–384.
9. Wandersman,C., and Yarmolinsky,M.B. (1977) *Virology* **77**, 386–400.
10. Appleyard,R.K. (1954) *Genetics* **39**, 440–452.
11. Sternberg,N. (1979) *Virology* **96**, 129–142.
12. Kondo,E., and Mitsuhashi, S. (1964) *J. Bacteriol.* **88**, 1266–1276.
13. Smith,H.W. (1972) *Nature New Biol.* **238**, 205–206.

14. Scott, J.R., Kropf, M.M. and Mendelson, L. (1977) *Virology* **76**, 39–46.
15. Fürste, J.P., Pansegrau, W., Frank, R., Blöcker, H., Scholz, P., Bagdasarian, M., and Lanka, E. (1986) *Gene* **48**, 119–131.
16. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene* **33**, 103–119.
17. Aiba, H., Adhya, S. and de Crombrughe, B. (1981) *J. Biol. Chem.* **256**, 11905–11910.
18. Thurston, C.F., Perry, C.R. and Pollard, J.W. (1988) In Walker, J.M. (ed.), *Methods in Molecular Biology—New Nucleic Acid Techniques*. Humana Press, Clifton, New Jersey, Vol. 4, pp 1–11.
19. Pollard, J.W., Perry, C.R. and Thurston, C.F. (1988) In Walker, J.M. (ed.), *Methods in Molecular Biology—New Nucleic Acid Techniques*. Humana Press, Clifton, New Jersey, Vol. 4, pp 13–25.
20. Feinberg, A.P., and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
21. Feinberg, A.P., and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–269.
22. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
23. Altman, S. (1989) *Advances in Enzymol.* **62**, 1–36.
24. Peck-Miller, K.A., and Altman, S. (1991) *J. Mol. Biol.* **221**, 1–5.