Proc. Natl. Acad. Sci. USA
Vol. 78, No. 4, pp. 2095–2099, April 1981
Biochemistry

The nucleotide sequence surrounding the replication terminus of R6K

(M13mp5 vector/gene 6 exonuclease of T7/DNA sequence determination by dideoxy method)

DEEPAK BASTIA*, JOSEPH GERMINO*, JORGE H. CROSA†, AND J. RAM‡

*Department of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710; †Department of Microbiology, University of Washington, Seattle, Washington 98165; and ‡Department of Microbiology, University of Alabama, Birmingham, Alabama 35294

Communicated by A. Dale Kaiser, December 18, 1980

ABSTRACT **The replication terminus of the plasmid R6K has been cloned into the single-stranded DNA phage vector M13mp5 and also into the plasmid vectors pBR313 and pBR322. By using single-stranded DNA templates prepared from the recombinant DNA clones, the sequence of 215 base pairs of DNA containing the replication terminus has been determined. The DNA sequence of the region of the terminus does not contain any 2-fold rotational symmetry. Therefore, folding of the DNA at the region of the terminus is unlikely to be a cause for replication termination. Interaction of a host-specified protein(s) with the sequence of the replication terminus is probably the basis of the mechanism of replication termination.**

In spite of extensive work, the molecular mechanism of termination of DNA replication still remains unclear. In efforts to elucidate the mechanism of termination of replication, analysis of the DNA sequences involved in replication termination and identification of proteins that interact with the replication terminus are of utmost importance.

The drug-resistance factor R6K which carries genes for ampicillin and streptomycin resistance (1) is a model system for this study because its replication is asymmetrically bidirectional. The asymmetry in the bidirectional replication is manifested in the confluence of the two moving forks of the replicating chromosome at a point that is not located at 180° from the replication origins (2, 3). The asymmetrically bidirectional replication strongly suggested the existence of a genetically determined terminus of replication.

In this paper we report the cloning of the replication terminus into the vector M13mp5, its precise localization in a 215-base-pair DNA fragment, and analysis of its nucleotide sequence. The nucleotide sequence of the replication terminus suggests a mechanism of termination that probably involves interaction of the terminus sequence of the plasmid with one or more proteins specified by the host genome.

MATERIAL AND METHODS

Bacterial, Plasmid, and phage Strains. The bacterial strain JC411 thy$^-$ was obtained from D. Helinski and the strain 71-18, from J. Messing (4). The plasmid vectors pBR313 and pBR322 (5) were obtained from H. Boyer through R. Curtiss III. The single-stranded DNA phage vector M13mp5, which contains a single *Hin*dIII site was a gift from J. Messing.

Chemicals, Enzymes, and Labeled Precursors. dNTPs (Sigma), dideoxynucleotide triphosphates (P-L Biochemicals), [α-^{32}P]dATP (100–300 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels; ICN), 5-bromo-4-chloroindolyl β-galactoside (Sigma) and isopropyl β-thiogalactoside; (Sigma) were obtained as indicated.

The restriction enzymes *Hin*dIII (6), *Hpa* II (7), *Alu* I (8), *Hinf* (9), *Hae* III and *Hae* II (10), and *Hha* (11) were purified according to published procedures. DNA polymerase I holoenzyme and the Klenow fragment of *Escherichia coli* were purchased from Boehringer Mannheim. Exonuclease III and *Bam*HI were purchased from New England BioLabs. T4 ligase was purified as described by Weiss (12).

Molecular Cloning. These experiments were conducted under the P1 level of containment. The fragment H_2 which contains the replication terminus was cloned into double-stranded M13mp5 DNA as described by Messing (4). The recombinant phage is called M13mp5H_2.

Preparation of DNA for Sequence Determination by the Exonuclease Method. Supercoiled RNA-free plasmid DNA was prepared by two cycles of dye/CsCl gradient centrifugation. The DNA was extracted with isopropanol, dialyzed, phenol extracted, and centrifuged in a 15–50% sucrose gradient as described (13). Supercoiled plasmid DNA (5 μg) was digested with *Bam*H to completion in a total volume of 20 μl. Then, 3 μl of 10-fold concentrated exonuclease buffer (10 mM $MgCl_2$/100 mM dithiothreitol/700 mM Tris, pH 7.5) was added followed by 5 units of exonuclease III, and the mixture was digested in sealed capillaries for 204 hr at 37°C (14). The digestions were terminated by heating to 68°C for 10 min. Digestion with T7 exonuclease was performed similarly except that 10 units of the exonuclease was used for 2 hr in a total reaction volume of 40 μl.

Single-Stranded M13 DNA. M13 phage from the supernatant of infected cultures were precipitated with 4% polyethylene glycol and purified in a CsCl gradient. Single-stranded phage DNA was purified by phenol extraction as described by Messing (4).

Digestion of Restriction Fragments with Exonuclease III. Restriction fragments from approximately 300 μg of plasmid DNA were extracted from the gels (15) and dissolved in 100 μl of 0.01 M Tris/1 mM EDTA, pH 8.00. Aliquots (5 μl) of the fragment solution were made to standard concentration exonuclease buffer (see above) and digested with 1 unit of exonuclease III for various periods of time, depending on the lengths of each fragment (e.g., 50–100 nucleotides for 3–5 min at 37°C or 200–300 nucleotides for 8–10 min at 37°C) in sealed capillary tubes. The reactions were terminated by heating to 68°C for 10 min.

DNA Sequence Determination. Each set of reactions was performed as described by Sanger *et al.* (16) with 5 μl of primer restriction fragment and 1 μg of either exonuclease III or T7 exonuclease-treated template. In the case of M13 clones, 1 μg of single-stranded DNA with 5 μl of primer was used. In those cases in which it was not necessary to cleave off the primer, the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

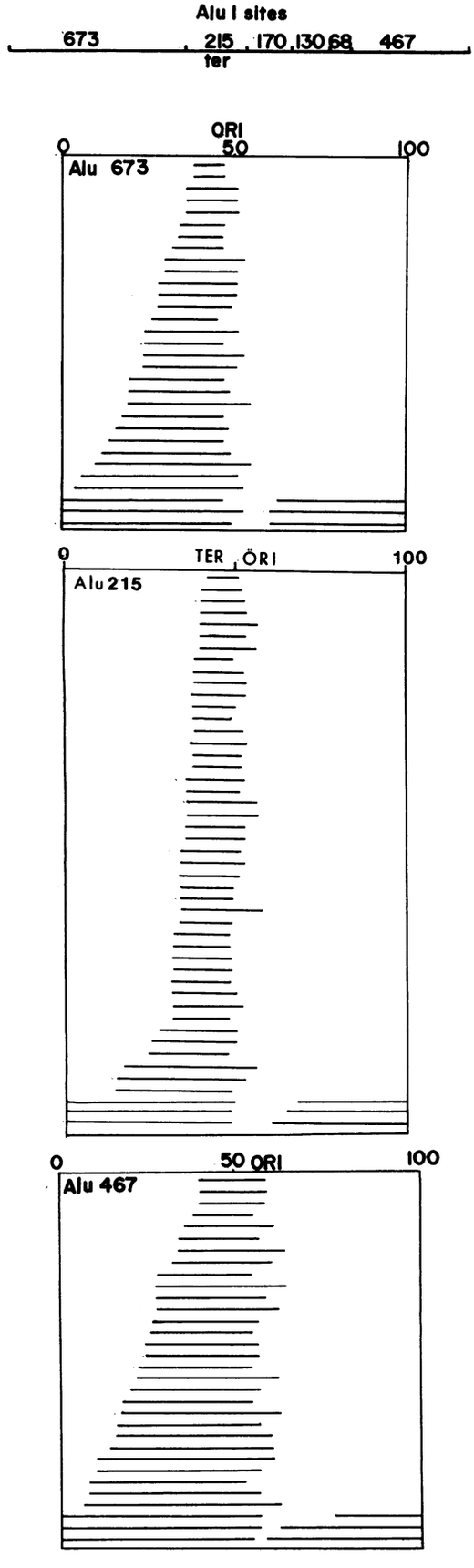

FIG. 1. Electron microscope analyses of replication intermediates of recombinant DNA subclones containing various *Alu* I fragments from the region of the replication terminus. The replication intermediates were linearized by digestion with *Eco*RI. The horizontal lines represent the replication loops. ORI, origin of replication; ter or TER, terminus of replication. (*Top*) Array of replication intermediates of the subclone Alu673. The location of the various *Alu* I fragments is shown in the restriction map shown at the bottom of the figure. Note the unidirectional replication of the chromosome of Alu673 with no apparent pause of the replication fork at any region of the chromosome. (*Middle*) Replication intermediates of Alu215. Note the sharp arrest of the moving fork at the TER sequence in 90% of the molecules examined. The results demonstrate that fragment Alu215 contains the replication terminus. (*Bottom*) Replication intermediates of the subclone Alu467. Like Alu673, Alu467 does not show arrest of the replication fork.

primer was also digested with exonuclease III as described above.

Computer Analysis of the Sequence Data. The DNA sequence data were analyzed by using a slight modification of the program published by Staden (17, 18).

Electron Microscopy of Replication Intermediates. Cultures of *E. coli* containing recombinant plasmid clones were grown in minimal essential medium to a cell density of 2×10^8 ml. Chloramphenicol was added to 170 μg/ml and the cultures were aerated for 5 hr at 37°C. Replication was stopped by adding 10 mM KCN/10 mM Tris/1 mM EDTA, pH 8.00 and rapidly cooling the culture to 0°C. The cultures were pelleted and the plasmid DNA was purified as described by Bastia (13). Replicating intermediates were purified by ethidium bromide/CsCl equilibrium centrifugations (2), linearized by digestion with *Eco*RI in 66 mM Tris/6 mM 2-mercaptoethanol/10 mM $MgCl_2$/200 mM NaCl. The DNA samples were prepared for electron microscopy as described by Bastia *et al.* (19).

RESULTS

Localization of the Replication Terminus. Experiments to be reported elsewhere had localized the replication terminus of R6K to a 2014-base-pair DNA fragment. The localization was based on electron microscopy of replication intermediates and isolation of open circular replication intermediates of mini R6K plasmids containing a discontinuity of both the plus and minus strands of newly replicated DNA. Upon filling the gaps with [α-^{32}P]dNTPs and *E. coli* DNA polymerase I, restriction enzyme cleavage, gel electrophoresis, and autoradiography of the gels revealed that the gap was located in a cluster of restriction fragment in the region of the replication terminus.

We attempted to localize the replication terminus more precisely within the 2014-base-pair fragment by subcloning various subfragments of the region. The order of *Alu* I-generated subfragments of the 2014-base-pair terminus region is shown in Fig. 1 Top. The fragments (Alu673, Alu215, Alu170, etc.) were cloned into the *Pvu* II site of the vector pBR322 by using the blunt-end joining activity of T4 DNA ligase.

Cultures of the subclones were grown separately, and replication intermediates from each subclone were purified by equilibrium centrifugation in dye/CsCl gradients, linearized by digestion with *Eco*RI, and prepared for electron microscopy. A random sample of replication intermediates was photographed, traced, and plotted in increasing order of replication. Arrays of replication intermediates of three representative subclones are shown in Fig. 1. The horizontal lines represent the replication loops. The replication intermediates of both Alu673 and Alu467 show unidirectional replication from the fixed replication origin, and there is no evidence of a pause of the moving fork at any region along the length of the chromosome; the replication terminates at or near the replication origin.

In contrast to Alu673 and Alu467, the replication intermediates of subclone Alu215 show a sharp arrest of the replication fork at a point approximately 500 base pairs from the replication origin. The point of arrest of the replication forks corresponds

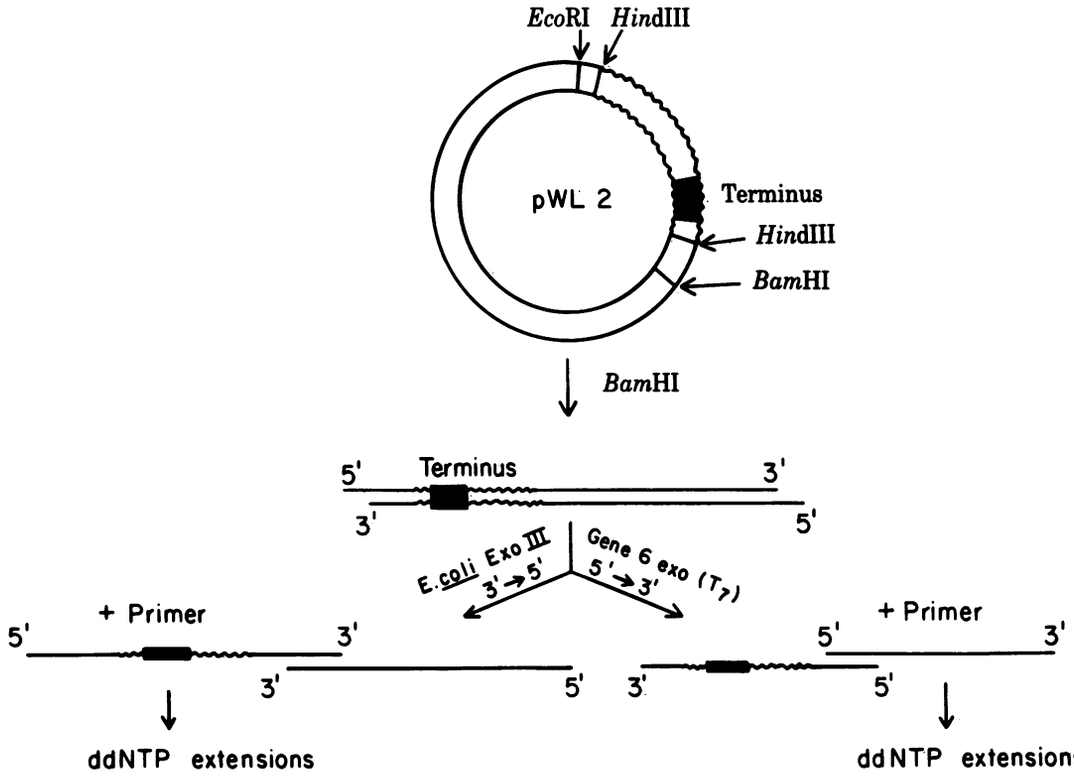

FIG. 2. Diagram of the procedure used to generate single strands from the region of the replication terminus. Approximately 5 μg of pWL2 DNA (the recombinant plasmid contains *Hin*dIII fragment 2 of R6K cloned into the single *Hin*dIII site of pBR313) was digested with 5 units of *Bam*HI for 1 hr at 37°C. The enzyme was inactivated by heating the reaction mixture to 68°C for 5 min. The linear double-stranded DNA thus generated was digested either with 5 units of exonuclease III or 5 units of T7 gene 6 exonuclease for 2–6 hr in exonuclease buffer in sealed 20 μl glass micropipettes. The exonuclease was inactivated by boiling the sealed micropipette in a water bath for 1 min. Then, 1 μg of exonuclease-treated DNA was mixed with 3 μl of primer DNA fragment (which was sometimes pretreated with exonuclease III).

to the location of the 215-base-pair segment cloned at the *Pvu* II site of the vector.

The arrest of the replication fork at the terminus sequence is transitory because the replication forks eventually traverse the replication terminus and finish a round of replication at the replication origin (Fig. 1 *Middle*).

This electron microscope analysis unambiguously localized the replication terminus sequence in the 215-base-pair *Alu* I fragment marked ter in Fig. 1 *Top*).

Strategies for DNA Sequence Analysis. We applied the chain termination method (17) of DNA sequence analysis to the 2014-base-pair region of DNA about the replication terminus of R6K. One essential prerequisite for this method is the availability of single-stranded DNA templates. We obtained single-stranded DNA templates by two procedures: (*i*) cloning and replication terminus into M13mp5, the single-stranded DNA phage vector (4), and (*ii*) digestion of linear pWL2 (a recombinant clone containing the replication terminus of R6K cloned into the *Hin*dIII site of the plasmid pBR313) DNA with either exonuclease III (6) or the gene 6-specified exonuclease of phage T7 (12).

Curiously, cloning into the M13mp5 vector always yielded the insertion of the replication terminus into the vector DNA in only one orientation. Therefore, only one of the two strands of the replication terminus region could be obtained by this procedure. Barnes (20) has described the use of M13 clones for analysis of nucleotide sequence by the dideoxy method (17). Although the M13 terminus clones yielded minicircles of various sizes, sequence analysis demonstrated that the deletions were always outside the region of the replication terminus.

We obtained confirmatory sequence data by digesting the plasmid pWL2 with *Bam*HI (Fig. 2) and separately digesting the linear double-stranded DNA with either exonuclease III or the gene 6 exonuclease of T7 (21). The best results were obtained by using single-stranded tails generated by T7 exonuclease as the template in the sequence reactions. The exonuclease III-generated templates sometimes yielded sequence gels with extra bands, presumably because some commercial samples of exonuclease III were contaminated with trace amounts of endonuclease.

A representative autoradiogram of a DNA sequence gel is shown in Fig. 3.

The sequence of the 215-base-pair region about the replication terminus is shown in Fig. 4. This region corresponds to the terminus region contained in the subclone Alu215.

Two features of the 215-base-pair terminus sequence are worth pointing out. The terminus region does not contain any 2-fold rotational symmetries. Therefore, the terminus is devoid of potential secondary structure. The other feature of the terminus sequence is the absence of a long open reading frame consisting of an initiator AUG or GUG codon and a ribosome binding site. For example, although AUG and GUG triplets are present at positions 702, 707, 732, and 766 of the strand shown in Fig. 4 and at positions 835, 823, 765, 753, and 734 in the complementary strand, one encounters chain-terminating triplets within 150 residues by reading from each of these initiating triplets. Also, the initiating triplets are not preceded by potential ribosome binding sequence. Therefore, the 215-base-pair terminus sequence is unlikely to code for a protein.

DISCUSSION

The results described in this paper shed light on a poorly understood step of DNA replication—the termination of replication—

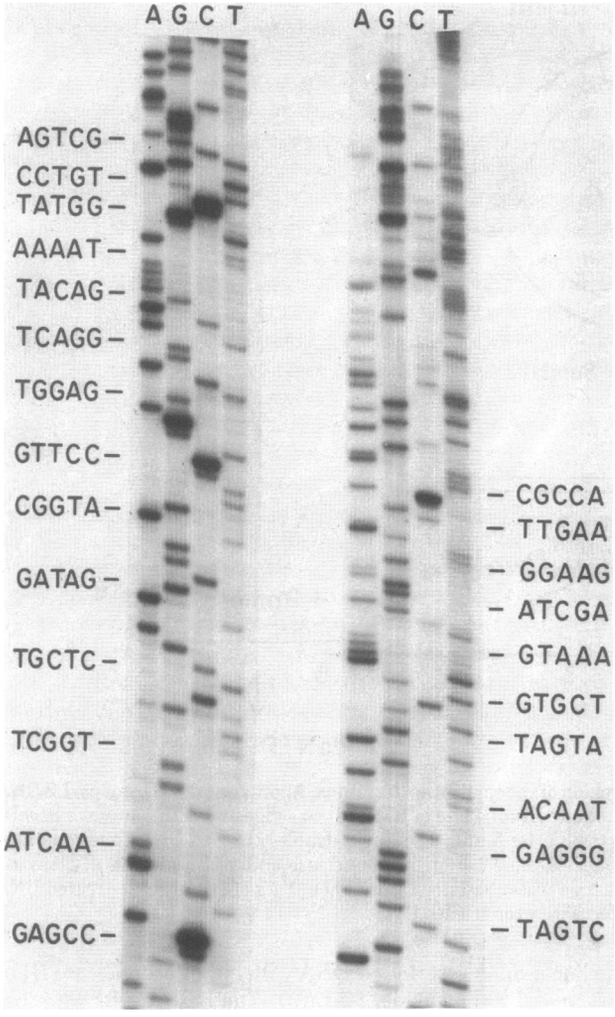

FIG. 3. Representative autoradiograms of DNA sequence gels. The template used in these two examples was single-stranded M13mp5H2 DNA.

by revealing the molecular anatomy of a sequence specific replication terminus.

The sequence-specific replication terminus of the plasmid R6K was unambiguously localized, by subcloning and electron microscopy, to the 215-base-pair *Alu* I fragment. Electron microscope analysis of recombinant DNA clones containing the 215-base-pair terminus sequence revealed that the terminus functions by transiently arresting the progression of a replication fork at or near the terminus sequence.

We have considered three hypotheses to explain the molecular mechanism of the termination of DNA replication.

(*i*) The terminus sequence sterically hinders the progression of replication forks because of its potential secondary structure.

(*ii*) The terminus sequence reversibly destabilizes one or more components of the replisome.

(*iii*) The terminus sequence has specific affinity for a protein which is normally not a constituent of the replisome and effects termination by binding to the terminus sequence.

The sequence analysis reported in this paper shows that the terminus does not contain any 2-fold rotational symmetry and therefore localized folding of the DNA at the region of the terminus is not a likely explanation for the arrest of the replication fork.

Further work is necessary to determine whether protein–DNA interaction causes termination of replication, to identify protein(s) having specific affinity for the terminus sequence, and to determine whether the terminus-binding protein is a part of the replisome.

The nucleotide sequence of the terminus reveals that the 215-base-pair sequence does not contain a significant open reading frame and therefore is unlikely to code for a protein. Further evidence that the terminus does not have a *trans*-acting function comes from *in vitro* DNA replication experiments (unpublished data) demonstrating that cell extracts of *E. coli* strains which do not harbor a resident terminus-sequence-containing plasmid are competent to effect termination of DNA replication *in vitro* at the specific replication terminus.

Although attempts to find specific termini of replication in simian virus 40 and λ (early stage of replication) have been unsuccessful (22, 23), the chromosome of *E. coli* has been unambiguously shown to contain a specific replication terminus located between the *rac* and *man* loci, in a genetically silent region (24, 25).

We thank Drs. J. Messing, W. Barnes, and N. Godson for their help and interest in our work. We are grateful to M. Shon and G. Cardineau for the purification of T7 exonuclease and to Deborah Story for typing this manuscript. This work was supported by Grant GM 24357 from the National Institute of General Medical Sciences to D.B. and by Grant PCM 75-14174-802 from the National Science Foundation to J.H.C.

1. Kontomichalou, P., Mitani, M. & Clowes, R. D. J. (1970) *Bacteriology* **104**, 34–44.
2. Crosa, J. H., Luttrop, L. K. & Falkow, S. (1976) *J. Bacteriol.* **129**, 454–466.
3. Lovett, M. A., Spraks, R. B. & Helinski, D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2905–2909.
4. Messing, J., Gronenborn, G., Muller-Hill, B. & Hoffschneider, P. H. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3642–3646.

```
         680          690          700          710          720          730
-AGCTCCAAG  CGAGCAGAGC  GAAACTACCA  GGATGATATG  ATCGAATCGT  TCGGCTACTG

         740          750          760          770          780          790
TCATGTTTAT  CCTCAAAAAG  CATTAAAAAT  TCCACGATGA  TTTAGTTACA  ACACACAAGA

         800          810          820          830          840          850
GATTACACTT  GCCCTTCCCT  CTGGCCTCCG  CACCATATTA GCCACACTCG  CAAGCCAAGC

         860          870          880          890
AGAGTCGGCT  TAAATTCTAT  TGAGTGTTGT  AACTACTAGC  T-
```

FIG. 4. Nucleotide sequence surrounding the replication terminus of R6K.

5. Bolivar, F., Rodriguez, R., Greene, P. J., Betlach, M., Heyneker, M. C., Boyer, H., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 75–93.
6. Smith, H. O. (1974) *Molecular Biology Series: DNA Replication and Biosynthesis*, ed. Wickner, R. B. (Dekker, New York), Vol. 7, pp. 71–85.
7. Sharp, P. A., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **14**, 4634–4638.
8. Roberts, R. J., Meyers, P. A., Morrison, A. & Murray, K. (1976) *J. Mol. Biol.* **102**, 157–165.
9. Subramanian, K. N., Weissman, S. M., Zain, B. S. & Robert, R. J. (1977) *J. Mol. Biol.* **110**, 297–317.
10. Roberts, R. J., Breitmeyer, J. B., Tabachnik, N. F. & Meyers, P. A. (1975) *J. Mol. Biol.* **91**, 121–123.
11. Roberts, R., Meyer, P. A., Morrison, A. & Murray, K. (1976) *J. Mol. Biol.* **103**, 199–208.
12. Weiss, R., Jaquemin-Sablon, A., Live, T. R., Fareed, G. & Richardson, C. C. (1968) *J. Biol. Chem.* **243**, 4543–4555.
13. Bastia, D. (1978) *J. Mol. Biol.* **124**, 601–639.
14. Smith, A. J. H. (1979) *Nucleic Acids Res.* **6**, 831–848.
15. Bastia, D. (1977) *Nucleic Acids Res.* **4**, 3123–3142.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
17. Staden, R. (1977) *Nucleic Acids Res.* **4**, 4037–4051.
18. Staden, R. (1978) *Nucleic Acids Res.* **5**, 1013–1015.
19. Bastia, D., Sueoka, N. & Cox, E. C. (1975) *J. Mol. Biol.* **98**, 305–320.
20. Barnes, W. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4281–4285.
21. Kerr, C. & Sadowski, P. D. (1972) *J. Biol. Chem.* **247**, 305–310.
22. Lai, C. S. & Nathans, D. (1975) *J. Mol. Biol.* **97**, 113–118.
23. Valenzuela, M. S., Friefelder, D. & Inman, R. (1976) *J. Mol. Biol.* **102**, 569–589.
24. Kuempel, P., Duerr, S. A. & Seeley, N. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3927–3931.
25. Louarn, J., Patte, D. & Louarn, J. M. (1977) *J. Mol. Biol.* **97**, 113–118.