Proc. Natl. Acad. Sci. USA
Vol. 76, No. 10, pp. 4946–4950, October 1979
Biochemistry

Fidelity of replication of phage φX174 DNA by DNA polymerase III holoenzyme: Spontaneous mutation by misincorporation

(editing/proofreading/error frequency/site-directed mutagenesis)

ALAN R. FERSHT*

Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305

Communicated by Arthur Kornberg, July 9, 1979

ABSTRACT DNA from φX174 is replicated *in vitro* with a fidelity similar to that found genetically. A mutation of TAG → TGG may be induced, however, by varying the concentrations of deoxynucleoside triphosphates, with a frequency proportional to [dGTP]²/[dATP]. This complex concentration dependence is consistent with the active participation of a proofreading mechanism that hydrolytically excises mismatched base pairs as they are formed. A simple kinetic analysis predicts that the frequency of misincorporation depends on the ratio of incorrect to correct deoxynucleoside triphosphates times the concentration of the *next* triphosphate in the sequence to be added. This suggests that spontaneous mutation by misincorporation depends crucially on the composition of the deoxynucleoside triphosphate pool.

The fidelity of replication of DNA is extremely high. The frequency of spontaneous point mutation in *Escherichia coli* is in the region of 10^{-8}–10^{-10} per base per replication (1), and the frequency in phages is within 2–3 orders of magnitude of this (2, 3). Such accuracy is very difficult to measure *in vitro* by use of synthetic templates and the misincorporation of radioactively labeled deoxynucleoside triphosphates because of trace levels of contaminants. A biological assay has recently been developed to overcome these difficulties and to avoid other artifacts; it uses the rate of reversion of the φX174 *am3* mutant during replication of its DNA *in vitro* (4, 5). [The mutant has a G → A transition at position 587 in the Sanger sequence (6), converting the TGG codon for tryptophan to the TAG terminator. The mutation has no effect on phage assembly or DNA replication, but just prevents cell lysis (7).] The initial study involved the partial replication of the viral (+) strand by using DNA polymerase I of *E. coli* and a restriction fragment as a primer (4). The resultant DNA was then transfected into spheroplasts, and the progeny phage or infective centers were assayed for revertants. However, in the absence of added mutagens, no revertants could be detected above the background reversion frequency [$(1.7 \pm 0.6) \times 10^{-6}$], and only a lower limit of 1/7700 could be placed on the accuracy of polymerase I. The second study used the reconstituted T4 replication system to replicate the replicative form (RF I) DNA of φX174 to produce progeny double-stranded RF DNA (5). Again, on transfection no revertants were detected above background, placing a lower limit of $5/10^7$ for the accuracy of the T4 polymerase.

In the present study, I apply this procedure to the RF DNA → single-stranded (ss) DNA step in the replication of φX174, using the reconstituted enzyme system of DNA polymerase III holoenzyme, *rep* and ssDNA-binding proteins from *E. coli*, and the phage-encoded gene *A* protein. In the presence of ATP and the four deoxynucleoside triphosphates, multiple single-stranded circular copies of the viral (+) strand are made by using the complementary strand of the RF I DNA as template (3, 8, 9). This is the stage *in vivo* at which RF DNA is multiplied to give progeny. The strategy is to vary the concentrations of the deoxynucleoside triphosphates during replication *in vitro* and to examine the concentration dependence of the error rate, if any.

MATERIALS AND METHODS

Materials. Deoxynucleoside triphosphates were obtained from Sigma. Thin-layer chromatography on Polygram polyethyleneimine (PEI) sheets [1 M HCO_2H/0.5 M LiCl or 4 M NaO_2CH (pH 3.4) (10)] showed that dATP, dTTP, and dCTP were greater than 99% triphosphate whereas dGTP contained 6.7% dGDP but no detectable amounts of other triphosphates (<0.1%). Oligo$(dT)_{12}$ and poly(dA) were obtained from Miles.

DNA polymerase III holoenzyme [3.5×10^5 units/mg (11)], ssDNA-binding protein [2.3×10^4 units/mg (12)], gene *A* protein [5.5×10^6 units/mg (13)], *rep* protein [4×10^7 units/mg (14)], and dUTPase [2×10^7 units/mg (15)] were kindly provided by U. Huebscher, R. Fuller, N. Arai, and J. Shlomai of this laboratory. φX174 and its DNA were the gifts of J. Kobori. φX174 RF I DNA was prepared by a modification of the Hirt procedure (16, 17).

Oligo$(dT)_{12}([^3H]dC)_{0.05}$ and oligo$(dT)_{12}([^3H]dT)_2$ were prepared as described by Brutlag and Kornberg (18).

Methods. φX174 RF I DNA (1.4×10^{-2} pmol) was replicated in a reaction mixture (25 μl) containing DNA polymerase III holoenzyme (300 ng, 105 units) ssDNA-binding protein (1.6 μg, 36.8 units), *rep* protein (25 ng, 1000 units), gene *A* protein (9 ng, 50 units), dUTPase (0.1 ng, 200 units), ATP (800 μM), dCTP (40 μM), various concentrations of dGTP, dATP, and [^{3}H]dTTP (900 cpm/pmol), Tris·HCl (55 mM, pH 7.5), dithiothreitol (6.5 mM), glycerol (20% vol/vol), $MgCl_2$ (10.8 mM), NaCl (76 mM), EDTA (0.26 mM), and bovine serum albumin (0.19 mg/ml). (The deoxynucleoside triphosphates were preincubated with the dUTPase for 2 min at room temperature.) After 1 hr at 30°C, duplicate samples were assayed for extent of incorporation of [^{3}H]dTMP by acid precipitation onto nitrocellulose filters while the remainder was extracted with phenol at 65°C (19), precipitated with ethanol and washed. The products were examined by electron microscopy by the Kleinschmidt formamide technique (20).

Spheroplasts from *E. coli* W3350 were prepared by the method of Guthrie and Sinsheimer (19). Transfection and infective center assays were performed as described by Benzinger (21), whereas progeny phage were produced by the method of Pagano and Hutchinson (22). Indicator bacteria were *E. coli*

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.

Abbreviations: RF I, covalently closed, circular superhelical (replicative form) φX DNA; ssDNA, covalently closed, single-stranded φX DNA

* Present address: Department of Chemistry, Imperial College of Science and Technology, London SW7, England.

C for wild type and E. coli CR (su^+) for *am3* (3, 7). Plaques were scored after 3–4 hr of incubation at 37°C. About 10^{-4}–10^{-3} infective centers were formed per DNA particle.

Infective Center Compared with Phage Release Assay. Transfections were performed under conditions where the number of infective centers produced is linear with DNA concentration (19). The number of progeny phage produced, however, is equal to the number of infective centers multiplied by the number of viable phage produced per center (≈100–200 in the present experiments). Thus, when mixtures of wild-type and mutant DNA, which may give different yields of progeny per infection, are analyzed, the infective center assay gives the better measure of the reversion frequency. It is seen later in Table 1, however, that the two procedures agree within a factor of 2.

RESULTS

The Product DNA. The incorporation of [^{3}H]dT into the DNA replicated *in vitro* corresponded on average to about five to eight circles produced per circle of input RF I DNA. The number of infective particles produced, determined by infective center assays using phage extracted samples of φX174 ssDNA for calibration, was about 50–60% of this. Examination by electron microscopy of the products formed at the lowest, intermediate, and highest concentrations of dGTP used (Table 1) showed the following compositions, respectively: single-stranded circles, 84, 87, and 85; single-stranded linear molecules, 11, 9, and 9; double-stranded relaxed circles, 4, 4, and 4; and double-stranded linear molecules, 0, 0, and 2. No remaining supercoiled RF I DNA was found. It is seen in Table 1 that the phage used for preparation of the RF I DNA contained about 2×10^{-6} revertants. After transfection of the DNA extracted from this phage, the progeny phage had a slightly increased proportion of revertants, about 4×10^{-6}. Both the RF I DNA and the DNA produced *in vitro* at 8.4 μM dGTP form about 3×10^{-6} revertant phage after transfection. Transfection of the DNA thus produces phage that is phenotypically very similar to the parent.

Increasing the concentration of dGTP relative to dATP to favor the A → G transition, however, caused a dramatic increase in the number of revertants. At 1010 μM dGTP and 10.0 μM dATP, the reversion frequency increased 2000-fold to 1/150, as measured by progeny phage release, and to 1/300, as determined by the number of infective centers. This increase clearly represents mutation *in vitro*. The data appear reliable. The results of the infective center assays paralleled those of the progeny release assays, despite being performed under vastly different ratios of DNA to spheroplasts. The individual assays were also consistent over a 3- to 10-fold range of dilution.

The presence of dGDP as an impurity did not perturb the experiments. Addition of dGDP to 400 μM did not alter the reversion frequency (Table 1).

Dependence of Reversion Frequency on Deoxynucleoside Triphosphate Concentration. The nonlinear relationship between the reversion frequencies and the concentrations of dGTP and dATP may be easily simplified. In Fig. 1, the logarithm of the reversion frequency is plotted against log [dGTP], with [dCTP] and [dTTP] held constant at 40 μM and [dATP] approximately constant at 9.5–12 μM. The slope of +2.0 shows that the frequency varies at [dGTP]^2. A similar plot (Fig. 2) against log [dATP] at 911 μM dGTP has a slope of −1.0, showing that the frequency varies as 1/[dATP]. Combining all the data (at 40 μM dTTP and dCTP), the reversion frequency calculated from the infective center assays is given by:

$$\text{Frequency} = 0.0333\,[\text{dGTP}]^2/[\text{dATP}]. \quad [1]$$

Calculated from the progeny phase assay, it is:

$$\text{Frequency} = 0.0659\,[\text{dGTP}]^2/[\text{dATP}]. \quad [2]$$

The fit is very good; the correlation coefficients are, respectively, 0.986 and 0.998.

Revertants Are Mainly Wild Type. Restoration of the lytic function with increasing [dGTP] indicates that the amber codon TAG is removed by the replacement of the T or A by G. The data in Table 1 and Fig. 2 show that dGTP competes with dATP. However, there is no competition between dGTP and

Table 1. Reversion frequencies of φX174 DNA synthesized *in vitro*

Concentration, μM			Wild-type revertants/ *am3* mutants,* × 10^6	
dGTP	dATP	dTTP	Infective centers	Progeny phage
				2.0 ± 0.2†
				3.9 ± 0.2‡
				3.0 ± 0.4§
8.4	9.5	40	—	2.9 ± 0.5
93	9.5	40	—	54 ± 16
202	12.0	40	156 ± 66	119 ± 5
404	12.0	40	650 ± 100	1110 ± 80
810	12.0	40	1400 ± 200	3280 ± 180
910	12.0	10	1510 ± 330	1930 ± 150
910	12.0	40	2430 ± 470	4670 ± 350
910	12.0	160	3700 ± 800	4480 ± 320
910¶	12.0	40	2630 ± 280	4690 ± 980
910	60.0	40	686 ± 72	941 ± 46
910	119.0	40	341 ± 52	564 ± 100
910	238.0	40	121 ± 22	107 ± 13
1010	10.0	40	3380 ± 410	6620 ± 170

* Mean ± SEM from two to five determinations. A total of 18,286 mutants and 12,855 revertants were scored, generally at least 100 for each set.
† Parent phage.
‡ ssDNA from parent phage.
§ RF I DNA template.
¶ 400 μM dGDP was also present.

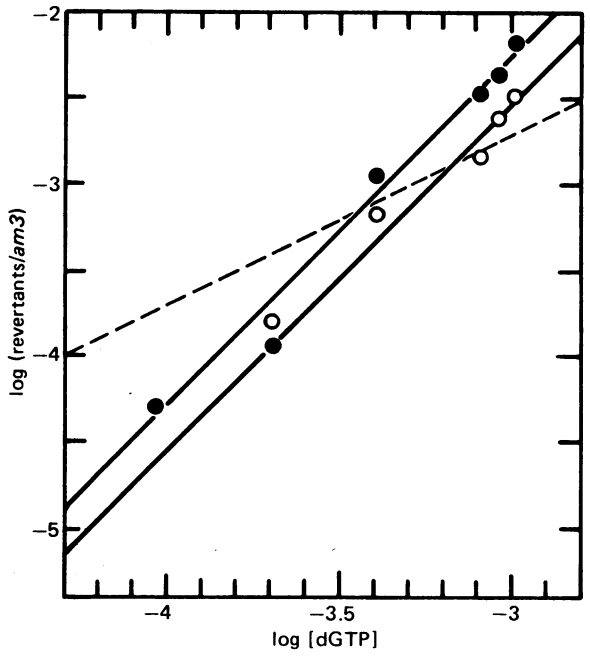


FIG. 1. Dependence of reversion frequency on [dGTP] at approximately constant [dATP] (9.5–12 μM). ●, Progeny phage assays; ○, infective center assays. Slope of solid lines = +2; slope of broken line = +1.

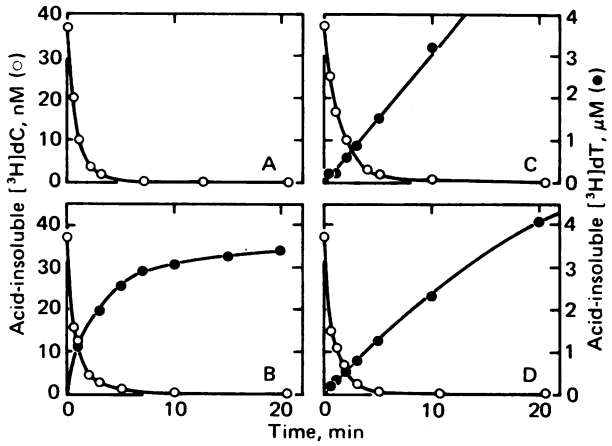


FIG. 3. 3′-5′ Exonuclease activity of DNA polymerase III holoenzyme as a proofreading function. 30°C, 75-μl reaction volumes. DNA polymerase III holoenzyme (600 ng, 200 units), $(dT)_{12}([^3H]dC)_{0.05}$ (665 pmol of nucleotide), Tris·HCl (50 mM, pH 7.5), dithiothreitol (5 mM), glycerol (11%, vol/vol), NaCl (12 mM), and bovine serum albumin (0.15 mg/ml). (*A*) 516 pmol of $(dA)_{600}$ and 0.8 μg of ssDNA-binding protein; (*B*) as *A* but plus 1.5 μmol of dTTP (○) and plus 1.5 μmol of [^{3}H]dTTP (2×10^7 cpm) (●); (*C*) as *B* but plus 6250 pmol of $(dA)_{300}$; (*D*) as *C* but without the DNA-binding protein. In all cases, the 3′-[^{3}H]dC was completely excised in about 10 min under polymerizing conditions.

dTTP; a 16-fold increase in [dTTP] at 911 μM dGTP leads to no decrease in reversion frequency. The codon must therefore be converted to TGG, the wild type.

At the higher concentrations of dGTP, however, the mutation frequencies suggest that multiple mutations may occur. This should not affect the calculations. It is suggested later (Eq. 6) that mutations are strongly biased to changing A-G to G-G. Because there are only about 170 of these pairs in φX174 that lead to amino acid changes, the majority of phage at even the highest [dGTP] should have no mutations induced. In support of this, the yield of viable phage was found to be constant over the entire range of concentrations. In any case, the effects of multiple mutation leading to nonviable phage should cancel out in the ratio of revertants to *am3* because both are derived mainly from the progeny.

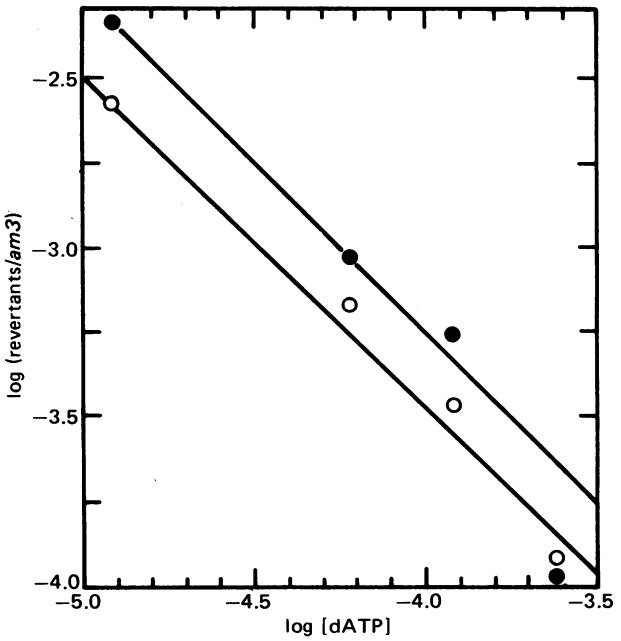


FIG. 2. Dependence of reversion frequency on [dATP] at constant [dGTP] (911 μM). ●, Progeny phage assays; ○, infective center assays. Slope of lines = −1.

3′-5′ Exonuclease Activity of DNA Polymerase III Holoenzyme Is a Proofreading Function. The 3′-5′ exonuclease activity of DNA polymerase I from *E. coli* appears to be a proofreading function for the removal of mismatched base pairs: mismatched bases at the 3′ terminus of synthetic primers are excised before elongation occurs (18). Although it has been shown that polymerase III from *E. coli* has the analogous 3′-5′ exonuclease activity (23, 24), it has not been shown that elongation occurs with the excision of the mismatches. The [^{3}H]dC at the 3′ terminus of $(dT)_{12}([^3H]dC)_{0.05}$ annealed to $(dA)_{600}$ was excised under a variety of polymerizing conditions in the presence of dTTP (Fig. 3). This is not caused by the 5′-3′ activity of the enzyme (23, 24) because, on replacing $(dT)_{12}([^3H]dC)_{0.05}$ by $(dT)_{12}([^3H]dT)_2$, the radioactivity was retained under polymerizing conditions. It is unlikely that the excision was caused by an impurity of polymerase I because the enzyme was prepared from a *polA*$^-$*polB*$^-$ mutant (11).

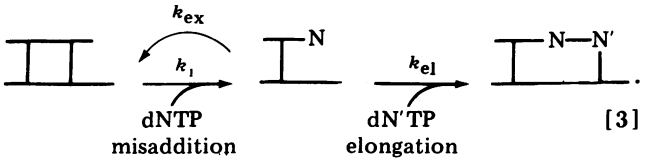
DISCUSSION

The replication of φX174 RF I DNA to ssDNA by the reconstituted enzyme system has several major advantages as a system in which to study fidelity and spontaneous mutation. (*i*) It uses the DNA polymerase III holoenzyme, believed to be the polymerase responsible for the major amount of replication of the *E. coli* chromosome. (*ii*) Only one strand of the template is copied and only one type of strand is produced, thus reducing complications in analyzing the effects of changing the concentrations of deoxynucleoside triphosphates. (*iii*) Multiple copies of the strand are produced, so that the progeny far outnumber the parent. (*iv*) The synthetic viral strands approximate the infectivity of authentic φX DNA.

Reversion Frequencies *In Vitro* Are Consistent with Those Found *In Vivo*. At low concentrations of dGTP, revertants are

at background level, but at high concentrations of dGTP to favor the A → G transition, revertants are detected with a frequency of 0.0333 $[dGTP]^2/[dATP]$. Extrapolation of this formula to concentrations of dATP and dGTP found in *E. coli* ($[dATP] \approx 56\ \mu M$, $[dGTP] \approx 28\ \mu M$, and $[dCTP] \approx [dTTP] \approx 50\ \mu M$; calculated from ref. 25 with a ratio of dry weight to wet weight = 0.25 and water space of *E. coli* = 72% (26)) predicts a reversion frequency of 0.5×10^{-6}, compared with an observed value of 1.7×10^{-6} (4).

Concentration Dependence of Reversion Frequency Is Consistent with a Proofreading Mechanism. If the 3′-5′ exonuclease activity of the polymerase acts during chain elongation as an editing mechanism to remove mismatched base pairs, then the simplest mechanism for misincorporation is the following:



[3]

The fraction incorporated is given by the product of the relative rates of addition of incorrect (dNTP) and correct (dN″TP) triphosphates to the chain multiplied by the probability that elongation, rather than excision, occurs. For two substrates competing for a common site, the former term is given by $(k_{cat}/K_M)_N[dNTP]/(k_{cat}/K_M)_{N''}[dN''TP]$ (27, 28), in which k_{cat} and K_M are the usual Michaelis–Menten quantities and N″ refers to dN″TP, etc., and the latter term is given by $k_{el}/(k_{el} + k_{ex})$.

It may be reasoned intuitively that the elongation rate, and hence the proofreading efficiency, must depend to some extent on the concentration of the *next* nucleotide (dN′TP) to be added. For example, at one extreme, when the next nucleotide is absent from the mixture, elongation will not occur and so errors are bound to be excised by the 3′-5′ exonuclease activity. It may be shown (see *Appendix*) that at low concentrations of the following nucleotide, the ratio of incorrect to correct nucleotide misincorporated (ν_{mis}) is given by:

$$\nu_{mis} = \alpha \frac{[dNTP][dN'TP]}{[dN''TP]}. \qquad [4]$$

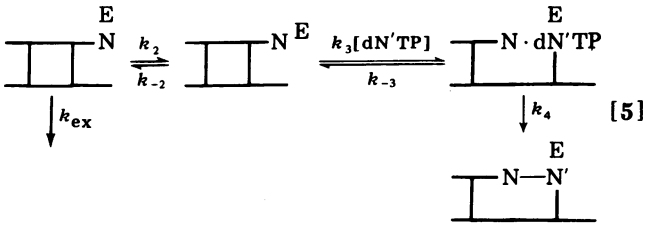
In the present study, both N and N′ = G and N″ = A. The observed concentration dependence of the reversion frequencies in this study is thus consistent with the 3′-5′ exonuclease activity being a proofreading function. The recent, independent observation that the misincorporation of 2-aminopurine into DNA with T4 polymerases increases with increasing concentrations of triphosphates is also consistent with this proposal (29).

Implications of Concentration Dependence of Error Frequency. If similar kinetic considerations apply *in vivo*, then spontaneous mutagenesis by nucleotide misinsertion depends crucially on the relative concentration of the correct and incorrect deoxynucleoside triphosphates and the absolute concentration of the following nucleotide to be inserted. This predicts that changes in the distribution of deoxynucleoside triphosphate pools will affect mutation rates. It is possible that some mutator genes function by changing the relative concentrations of nucleotides. The same criteria apply to the measurement of error rates *in vitro*. Eq. 4 and those below make predictions that can be tested in other systems and with other codons and that will aid in experimental design.

APPENDIX

Concentration dependence of misincorporation kinetics

To illustrate the intuitive arguments above, consider one of the simplest mechanisms for the elongation/excision step (Eq. 5). The enzyme, having misinserted N instead of N″, may either



[5]

move to the new primer terminus (where binding of dN′TP and elongation occur) or excise N while at the previous terminus. For simplicity, it is assumed that the frequency of excision of correctly paired nucleotides is negligibly small. The ratio of excision to elongation may be calculated either by application of the steady-state assumption or by the method of Cleland (30). By the logic and notation described for Eq. 3, it may be shown that:

$$\nu_{mis} = \frac{\beta k_2 k_3 k_4 [dN'TP]}{k_3 k_4 [dN'TP](k_2 + k_{ex}) + k_{-2} k_{ex}(k_{-3} + k_4)}, \qquad [6]$$

in which $\beta = (k_{cat}/K_M)_N[dNTP]/(k_{cat}/K_M)_{N''}[dN''TP]$.

The term in [dN′TP] follows saturation kinetics: at low [dN′TP], ν_{mis} is linear in [dN′TP] and reduces to Eq. 4; at high [dN′TP], ν_{mis} becomes independent of [dN′TP] and tends to $\beta k_2/(k_2 + k_{ex})$. In the present study, the misincorporation of dGTP is still in the linear region of the curve at 1 mM, whereas the misincorporation of the unnatural 2-aminopurine (which is misinserted at a relatively high frequency) saturates at lower concentrations (29).

Mechanisms more explicit and complex than Eq. 5 (e.g., in refs. 29 and 31) may also be shown to give misincorporation kinetics similar to Eq. 6.

The work reported in this paper was undertaken during the tenure of an American Cancer Society–Eleanor Roosevelt–International Cancer Fellowship awarded by the International Union Against Cancer. I am deeply indebted to Dr. Arthur Kornberg in whose laboratory this study was performed.

1. Fowler, R. G., Degnen, G. E. & Cox, E. C. (1974) *Mol. Gen. Genet.* **133**, 179–191.
2. Drake, J. W. (1970) *The Molecular Basis of Mutation* (Holden-Day, San Francisco).
3. Denhardt, D. T. & Silver, R. B. (1966) *Virology* **30**, 10–19.
4. Weymouth, L. A. & Loeb, L. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1924–1928.
5. Lin, C. C., Burke, R. L., Hibner, U., Barry, J. & Alberts, B. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 469–487.
6. Sanger, F., Coulson, A. R., Friedmann, T., Air, G. M., Barrell, B. G., Brown, N. L., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M. & Smith, M. (1978) *J. Mol. Biol.* **125**, 225–246.
7. Hutchison, C. A. & Sinsheimer, R. L. (1966) *J. Mol. Biol.* **18**, 429–447.
8. Eisenberg, S., Scott, J. F. & Kornberg, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3151–3155.
9. Eisenberg, S., Scott, J. F. & Kornberg, A. (1978) in *The Single Stranded DNA Phages*, eds. Denhardt, D. T., Dressler, D. & Ray, D. S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 287–302.
10. Randerath, K. & Randerath, E. (1964) *J. Chromatog.* **16**, 111–125.

11. McHenry, C. & Kornberg, A. (1977) *J. Biol. Chem.* **252,** 6478–6484.
12. Weiner, J. H., Bertsch, L. L. & Kornberg, A. (1975) *J. Biol. Chem.* **250,** 1972–1980.
13. Scott, J. F., Eisenberg, S., Bertsch, L. L. & Kornberg, A. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 193–197.
14. Scott, J. F. & Kornberg, A. (1978) *J. Biol. Chem.* **253,** 3292–3297.
15. Shlomai, J. & Kornberg, A. (1978) *J. Biol. Chem.* **253,** 3305–3312.
16. Hirt, B. (1967) *J. Mol. Biol.* **26,** 365–369.
17. Eisenberg, S., Harbers, B., Hours, C. & Denhardt, D. T. (1975) *J. Mol. Biol.* **99,** 107–123.
18. Brutlag, D. & Kornberg, A. (1972) *J. Biol. Chem.* **247,** 241–248.
19. Guthrie, G. D. & Sinsheimer, R. L. (1963) *Biochim. Biophys. Acta* **72,** 290–297.
20. Simon, M., Davis, R. & Davidson, N. (1971) *Methods Enzymol.* **21,** 413–428.
21. Benzinger, R. (1978) *Microbiol. Rev.* **42,** 194–236.
22. Pagano, J. S. & Hutchison, C. A. (1971) *Methods Virol.* **5,** 79–123.
23. Livingston, D. & Richardson, C. (1975) *J. Biol. Chem.* **254,** 1748–1753.
24. McHenry, C. S. & Crow, W. (1979) *J. Biol. Chem.* **254,** 1748–1753.
25. Munch-Peterson, A. (1970) *Eur. J. Biochem.* **15,** 191–202.
26. Roberts, R. B., Cowie, D. B., Abelson, P. H., Bollon, E. T. & Britten, R. J. (1957) *Studies of Biosynthesis in Escherichia coli* (Carnegie Inst. Washington, Washington, DC), pp. 58–60.
27. Fersht, A. R. (1974) *Proc. R. Soc. Lond. Ser. B* **187,** 397–407.
28. Fersht, A. R. (1977) *Enzyme Structure and Mechanism* (Freeman, San Francisco), pp. 96–97.
29. Clayton, L. K., Goodman, M. F., Branscomb, E. W. & Galas, D. J. (1979) *J. Biol. Chem.* **254,** 1902–1912.
30. Cleland, W. W. (1974) *Biochemistry* **14,** 3220–3224.
31. Galas, D. J. & Branscomb, E. W. (1978) *J. Mol. Biol.* **124,** 653–687.