

A DNA polymerase α pause site is a hot spot for nucleotide misinsertion

(mutagenesis/fidelity)

MICHAEL FRY*[†] AND LAWRENCE A. LOEB[‡]

*Bruce Rappaport School of Medicine, Unit of Biochemistry, Technion–Israel Institute of Technology, P.O. Box 9694, Haifa, Israel; and [†]The Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology SM-30, University of Washington, Seattle, WA 98195

Communicated by Daniel Mazia, October 9, 1991

ABSTRACT In this study we examined whether the arrest of DNA polymerase α (pol α)-catalyzed DNA synthesis at template pause sites entails terminal nucleotide misincorporation. An approach was developed to identify the 3'-terminal nucleotide in nascent DNA chains that accumulate at pause sites. A radioactive 5'-end-labeled primer was annealed to a bacteriophage M13mp2 single-stranded DNA template and elongated by pol α . Individual DNA chains that were accumulated at pause sites were resolved by sequencing gel electrophoresis, isolated, and purified. These DNA chains were elongated by pol α by using four annealed synthetic DNA templates, each of which contained a different nucleotide at the position opposite the 3' terminus of the arrested chain. Owing to the high preference of pol α for matched-over-mismatched primer termini, only those templates that contain a nucleotide that is complementary to the 3' terminus of the isolated pause-site chain are copied. Electrophoresis of product DNA showed the extent of copying of each template and thus identified the 3'-terminal nucleotide of the pause-site chains. We found that product DNA chains terminate with a noncomplementary 3'-terminal nucleotide opposite pause sites within the sequence 3'-d(AAAA)-5' at positions 6272–6269 of the M13mp2 genome. pol α catalyzed misincorporation of dG or dA into the 3' terminus of nascent chains opposite two of the M13mp2 template dA residues. A similar analysis of a different pause site did not reveal significant misincorporation opposite template dC. These results suggest that some but not all sites at which pol α pauses may constitute loci of mutagenesis.

Spontaneous mutations are not uniformly distributed over the genome. Benzer (1) first demonstrated clustering of spontaneous mutations at specific positions in the rII system of bacteriophage T4. Detailed analysis of spontaneous mutations in the *lacI* gene in a mismatch repair-defective strain of *Escherichia coli* indicated that errors due to DNA replication are predominantly single base changes and that they are concentrated at specific sites along the DNA at frequencies that exceeded by 10- to 100-fold the frequency of mutations at other sites (2). The distribution of single base changes in DNA of mammalian cells also appears to be highly uneven; clustering of point mutations at selected positions was observed in genes such as *aprt* (3), *ras* (4), and *p53* (5). It has been suggested that the interaction of DNA polymerases with specific DNA sequence contexts leads to an augmented error rate at such loci. This proposition is supported by the observation that different eukaryotic DNA polymerases generate mutational hot spots while replicating *in vitro* the M13 *lacZ α* gene (6, 7). However, no mechanism has yet been invoked to explain how DNA polymerases produce mutations at an increased frequency at specific loci in DNA.

A general positive correlation between the processivity of DNA polymerases and their accuracy has been noted (8). The overall processivity of DNA polymerases is determined, *inter alia*, by their proclivity to pause along the template in the course of DNA synthesis. We examined, therefore, whether pause sites produced along M13mp2 DNA by DNA polymerase α (pol α), a major DNA replication and repair enzyme of mammalian cells (9), are associated with misincorporation of the 3'-terminal nucleotide into the growing DNA chain. Here we report that DNA chains that accumulate at some but not all pause sites contain a noncomplementary 3'-terminal nucleotide. These results suggest that some template barriers for DNA polymerase might constitute loci of increased mutagenesis.

MATERIALS AND METHODS

Isolation of DNA Chains That Accumulate at Pause Sites. The 16-mer synthetic primer 5'-d(GCTGCGCAACTGT-TGG)-3' (Operon Technologies, Alameda, CA) was labeled at its 5' terminus by using [γ -³²P]ATP (10) and hybridized directly to circular single-stranded M13mp2 DNA at a ratio of 1.8:1.0 primer/template molecules (11). The labeled primer, which complements nucleotides 6376–6262 of the M13mp2 genome, was extended *in vitro* by calf thymus DNA pol α -primase that was purified by immunoaffinity chromatography, and its units of activity were defined as described by Perrino and Loeb (12). Unless otherwise stated, DNA synthesis was conducted for 30 min at 37°C in a reaction mixture that contained in a final volume of 15 μ l: 20.0 mM Hepes buffer (pH 7.8), 1.0 mM dithiothreitol, 3.0 mM MgCl₂, 20.0 μ M of each of the four dNTPs, 0.15 unit of pol α , and 70 ng of primed M13mp2 DNA. Primer extension was terminated by the addition of EDTA (pH 8.0) to a final concentration of 6.0 mM, the volume was increased to 100 μ l with H₂O, and the mixture was centrifuged through a Sephadex G-50 minicolumn to remove salt and unincorporated [γ -³²P]ATP (10). The DNA was dried and heat-denatured and then was resolved by electrophoresis through an 8% sequencing polyacrylamide gel (11). The copied DNA was analyzed alongside a sequence ladder of M13mp2 DNA [prepared with a Sequenase kit (United States Biochemical)], and the undried gel was exposed to Kodak x-ray film to determine the location and nucleotide sequence of bands at pause sites. The gel was aligned on top of the autoradiogram and well-separated bands were cut out of the gel. The precision and extent of removal of each band were verified by a second autoradiography of the gel after excision. Finely minced gel slices that were pooled from three to six identical lanes were suspended in 100 μ l of 20 mM Tris-HCl, pH 8.0/1.0 mM EDTA/300 mM NaCl,

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: pol α , DNA polymerase α ; pol β , DNA polymerase β ; pol I, *Escherichia coli* DNA polymerase I; T7 pol, bacteriophage T7 DNA polymerase.

[†]To whom reprint requests should be addressed.

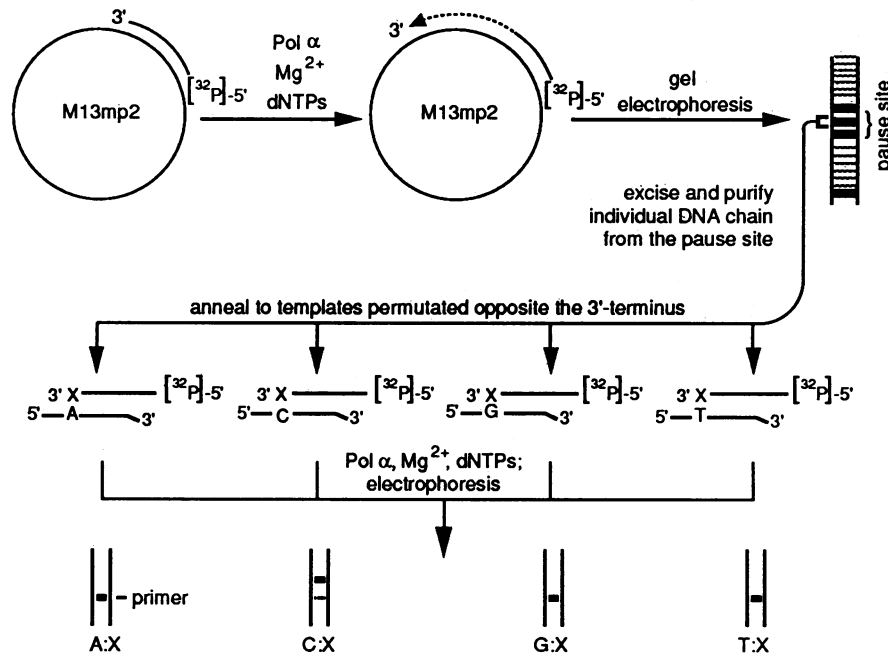


FIG. 2. Experimental strategy for the identification of the 3'-terminal nucleotide in nascent DNA chains that accumulate at pause sites. Primed M13mp2 DNA is copied by pol α , and product DNA chains are resolved by electrophoresis. DNA chains that accumulate at individual pause-site bands are excised from the gel, extracted, purified, and annealed to four 18-nucleotide-long complementary templates that differ by the nucleotide positioned opposite the 3' terminus of the isolated DNA chain. The annealed chains are extended by pol α , and the extent of synthesis is assessed by gel electrophoresis. In this example, the 3'-terminal nucleotide complements a template dC residue.

proportion of primer extended with each of the four templates is assessed by sequencing gel electrophoresis. A template-matching 3' terminus of the primer is extended by pol α at a rate that is 10^3 – 10^6 higher than the rate of extension of a mismatched terminal pair (12, 13). Hence, the 3'-terminal nucleotide of the isolated pause-site chain is defined by the complementing matching template nucleotide that allows its extension (Fig. 2).

Pause Site II Is a Locus of 3'-Terminal Nucleotide Misinsertion. DNA chains synthesized by pol α accumulated at each of the four consecutive dA template bases that constituted pause site II (Fig. 1). We first determined whether or not the terminated DNA chains opposite each of the template dA residues contained a matching 3'-dTTP. The 3' termini of product DNA chains, 107–110 nucleotides long, that accumulated opposite each residue of the template sequence 5'-d(AAAA)-3' of pause site II were designated N₁, N₂, N₃, and N₄, respectively. We chose a d(C)₃ template stretch after d(A)₄ since it was efficiently copied to completion. Individual DNA chains were purified, and each was annealed to a different complementary oligonucleotide template that contained a dA residue opposite the N₁–N₄ 3' terminus of each of the chains. pol α was added to elongate each of the four pause-site II chains as well as to extend a synthetic oligomeric primer with the sequence 5'-d(TTTT)-3' at its 3' end that was hybridized to the same template. After incubation for 30 min at 37°C, extension was assessed by the addition of one to three nucleotides to the primer. As seen in Fig. 3, the synthetic primer was increasingly elongated in the presence of 0.002 to 0.1 mM dNTPs. Note, however, that the copying of the ultimate dC template residue was inefficient (Fig. 3). In contrast to the control primer, none of the four isolated pause-site II chains was extended at any dNTP concentration (Fig. 3). Further, no elongation of the isolated pause-site II DNA chains was detected even after incubation for 90 min at 37°C (data not presented). That all of the four primer-templates were usable by DNA polymerases is reflected by their efficient copying with T7 pol and by the large fragment of *E. coli* pol I (results not shown). Both of these enzymes can

extend mismatched termini and/or excise mismatched 3' termini by their 3' → 5' exonuclease and subsequently extend the processed primer (14). Thus, the failure of pol α to extend all four pause-site II chains may be due to the presence of a nucleotide other than dT at their 3' terminus.

To identify the N₃ 3'-terminal nucleotide in the purified 109-residue-long pause-site II DNA chain, we hybridized the purified chain to four separate complementary oligomer templates, each of which contained a dA, dC, dG, or dT residue opposite the 3' terminus of the pause-site primer. The degree of primer extension by pol α with each template was

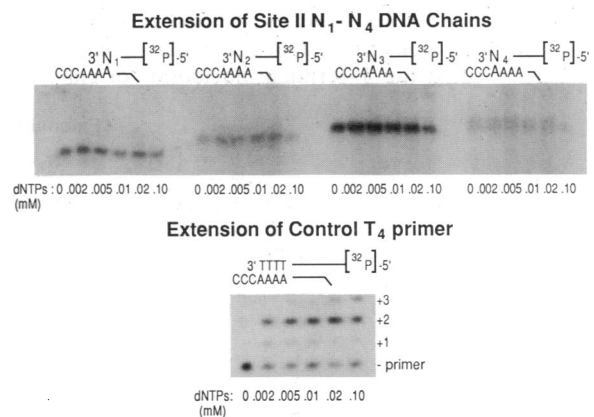


FIG. 3. The 3'-terminal nucleotide of each of the four DNA chains at pause site II is not extended by pol α when each is opposite a template dA. DNA chains, 5'-³²P-labeled, that terminate at their 3' end with nucleotides designated N₁, N₂, N₃, and N₄ were purified from the four corresponding gel bands of pause site II and annealed to complementary oligonucleotides that contained a dA residue opposite each N_n 3' terminus. Control 5'-³²P-labeled T₄ primer that terminated with four consecutive dT residues at its 3' end was annealed to the same template. pol α was used to extend each of the primers with the indicated concentrations of dNTP. Shown are autoradiograms of 8% polyacrylamide gels used to separate the extended DNA chains.

determined by gel electrophoresis (see Fig. 2). A typical result of one of four independent similar experiments is presented in Fig. 4. Only after incubation of 10 and 20 min at 37°C, a minority of the N₃ 3'-termini of pause-site II chain was elongated when positioned opposite template dA. In other, similar experiments, no extension was detected with this template, even after incubation for 30 min at 37°C (data not shown). From these data we estimate that only 10–30% of the purified chains contained dT in their 3' terminus. In all the experiments, no elongation was detected when the N₃ 3' termini were paired with template dG, indicating that dC is not the terminal nucleotide (Fig. 4). More than 50% of the N₃ termini were extended, however, when they were positioned opposite template dC residue, and some were elongated when paired with template dT residue (Fig. 4). That all four isolated pause-site II DNA chains were utilizable as primers when hybridized to any of the four oligomer templates was indicated by their extension with the large fragment of pol I (Fig. 4). In clear contrast to the isolated pause-site II N₃ chains, a synthetic control primer was extended by pol α when its dT 3' terminus was positioned opposite a dA template residue. However, in accord with previous studies (14, 15), pol α failed to detectably extend this primer when its 3'-terminal dT residue was positioned opposite a mismatched dC, dG, or dT template nucleotide (Fig. 4). Hence, the arrest of pol α at the third residue of pause-site II involves terminal incorporation of dGMP or, less frequently, dAMP opposite template dA.

Terminal misincorporation was also observed opposite the fourth dA template residue at pause site II. An N₄ synthetic control primer was extended by pol α when the dT residue at its 3' terminus was correctly paired with template dA but not when it was positioned opposite a mismatched dC, dG, or dT residue (Fig. 5). By clear contrast, the DNA chain terminating opposite the fourth dA residue at pause site II was not extended significantly by pol α when its 3'-terminal nucleotide was positioned opposite template dA, dG, or dT residue. However, this chain was efficiently extended when its 3'-terminal N₄ nucleotide was paired with template dC (Fig. 5). The failure of pol α to extend dA-, dG-, or dT-containing templates is not due to defects in these primer templates, as demonstrated by their efficient utilization by T7 pol (Fig. 5).

Not Every Pause Site Is a Locus for Nucleotide Misinsertion. To test whether or not misincorporation occurs at the 3' termini of every DNA chain that accumulated at any pause site, we identified the 3'-terminal nucleotide of chains that

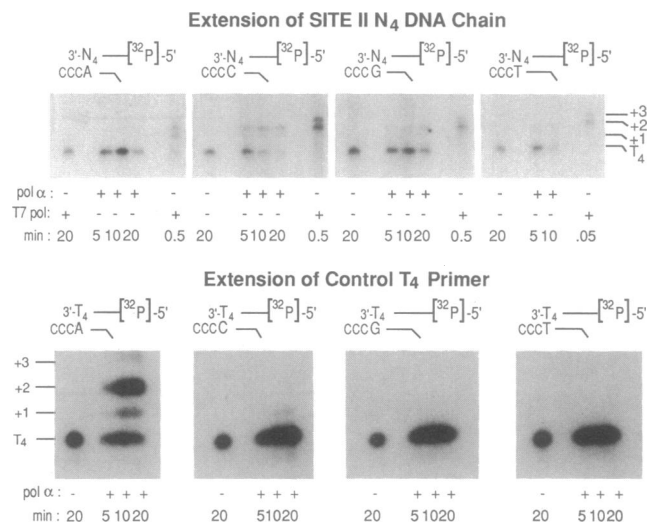


FIG. 5. Pause-site II chain N₄ terminus is extended by pol α mainly when paired with template dC. DNA chain, 5'-³²P-labeled, that terminated with a 3'-N₄ nucleotide was purified from the corresponding gel band of pause site II and hybridized to each of four oligomer templates that contained a dA, dC, dG, or dT residue opposite the N₄ terminus. In parallel, control 5'-³²P-labeled T₄ primer that terminated with a 3'-dT residue was annealed to the same four templates. Extension and electrophoresis are as indicated in Fig. 4.

were terminated opposite position dC residue 6345 at pause site I (Fig. 1). These 34-nucleotide-long DNA chains were purified, and the identity of their 3' terminus was established as described for the N₃ and N₄ termini of pause-site II chains. The 3' terminus of the pause-site I DNA chain was extended when it was positioned opposite an oligomeric template dC residue but not when it was paired with dA, dG, or dT (data not shown). These chains were thus correctly terminated with a dG residue and, hence, pausing by pol α is not invariably linked to nucleotide misinsertion.

DISCUSSION

Synthesis *in vitro* of DNA by isolated viral, bacterial, and eukaryotic DNA polymerases is commonly characterized by an interrupted progression of these enzymes along the template. Many purified DNA polymerases halt or detach from the template in the course of synthesis at discrete sites. This

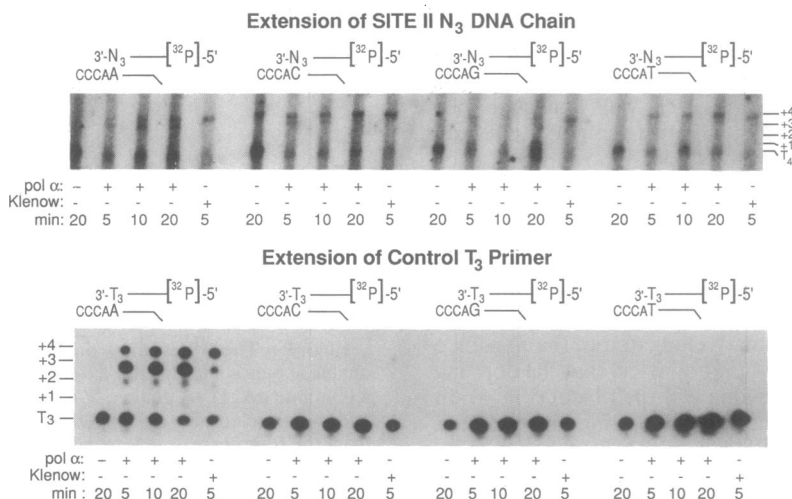


FIG. 4. Pause-site II chain N₃ terminus is extended by pol α mainly when paired with template dC. DNA chain, 5'-³²P-labeled, that terminated with a 3' N₃ nucleotide was purified (pause site II) and hybridized to each of four oligomer templates that contained a dA, dC, dG, or dT residue opposite the N₃ terminus. Control 5'-³²P-labeled T₃ primer that terminated with a 3'-dT residue was annealed to the same four templates. pol α was used to extend each of the primers, and 8% polyacrylamide gels were used to separate the extended DNA chains.

slowing or arrest of DNA synthesis causes bands of uncommonly high density on DNA sequencing gels. The location of polymerase pausing sites and their intensity are determined by their nucleotide sequence and, presumably, by the local structure of the DNA (15–20). Pausing is also affected by the association of binding proteins with the DNA (17, 21–24), by the type of polymerase used (11, 16, 19, 20–22), and by its combination with auxiliary proteins (21, 25–28).

Although the underlying mechanisms for the pausing of DNA polymerases at defined loci along the DNA are still largely unknown, it is generally accepted that the frequency of enzyme dissociation from the template and reinitiation of synthesis are increased at these sites. In the case of DNA polymerases devoid of 3' → 5' exonuclease, dissociation of the enzyme from the template could be caused by the incorporation of a noncomplementary nucleotide. The incorporation of a mismatched 3'-terminal nucleotide into the growing DNA chain results in diminished association of the enzyme with the template–primer (12, 13). Conversely, repeated dissociation and reassociation of a polymerase with the template at an arrest site might entail a high frequency of incorporation of a noncomplementary nucleotide. In fact, Hopfield (29) proposed that the addition of the first nucleotide to a primer terminus could constitute an error-prone step. The possible link between the pausing of DNA polymerases and decrease in fidelity has been indirectly addressed. Comparative measurements by the M13mp2 *lacZα* forward mutation assay of the fidelity of different DNA polymerases established a rough positive correlation between their processivity and their degree of accuracy (8). Bebenek *et al.* (30) compared the locations of pause sites for human immunodeficiency virus (HIV) and avian myeloblastosis virus reverse transcriptases with the distribution of errors in incorporation along the M13mp2 *lacZα* gene. Although no correlation was found between positions of pause sites and single-base changes, loci of pause sites generally coincided with single-base frameshift errors (30). It should be noted, however, that HIV reverse transcriptase can extend mismatched termini at a rate 50-fold greater than pol α (31); thus, the linking of pause sites with terminal misincorporation by HIV polymerase may be much less stringent than that for pol α .

The present study focused on the detection of possible misincorporation events that may occur at two selected pause sites of pol α along the M13mp2 *lacZα* region. Our assay utilizes the well-established high selectivity of pol α for extension of matched over mismatched base pairs at the 3' terminus (12, 13). Reaffirming this selectivity, we show here that pol α elongates efficiently a template-matching 3' terminus of the primer, but it fails to utilize to a significant extent mismatched termini (Figs. 4 and 5). We utilized this high degree of selectivity of pol α to identify unknown 3' termini of three nascent DNA chains that accumulated at pause sites. Whereas no misincorporation was detected opposite a template dC residue at pause site I (see *Results*), gross misinsertion of mainly dG and of some dA occurred opposite two template dA residues at pause site II (Figs. 4 and 5). Additional results suggest that the two remaining nascent DNA chains that accumulated at site II were also terminated by a misinserted nucleotide (Fig. 3). It is notable that spectra of mutations produced by pol α along the M13mp2 *lacZα* DNA stretch show no clustering of mutations at the pause-site II region (6). This discrepancy between terminal misincorporation and mutations is explained by the absence of a mutant phenotype in 8 of the 12 possible substitutions (T. A. Kunkel, personal communication).

The high preference of pol α for a matching base pair at the 3' terminus of the primer (refs. 12 and 13 and results in this paper) raises the intriguing possibility that some of the pause sites generated by this enzyme may be hot spots for nucle-

otide misincorporation rather than physical barriers that only block DNA synthesis. Hence, if a local structure within the template induces misincorporation, product DNA chains will be terminated with a mismatched nucleotide, pol α will fail to extend them, and nascent chains will be accumulated. Loci such as pause site II may thus directly represent foci of misincorporation. Alternatively, however, physical blocking of the polymerase and its detachment from the template and reattachment may increase misincorporation.

The results of this study demonstrate that some pause sites for pol α might constitute loci of increased misincorporation. This observation offers the tantalizing possibility that *in vivo* mutational hot spots in some cases may be generated during the pausing of a replicating DNA polymerase. Such a possibility is testable since genes with well-defined highly mutable nucleotide clusters such as *ras* (4) or *p53* (5) can be copied *in vitro*, and the locations of pausing and misinsertion can be directly evaluated.

M.F. is an American Cancer Society, Eleanor Roosevelt International Cancer Research Fellow 1990–1991. This study was supported by grants to M.F. from the United States–Israel Binational Fund, the Fund for Basic Research administered by the Israel Academy of Science and Humanities, and the Israel Cancer Association and by National Cancer Institute Outstanding Investigator Grant R35-CA-39903 to L.A.L.

1. Benzer, S. (1961) *Proc. Natl. Acad. Sci. USA* **47**, 403–416.
2. Schaaper, R. M. & Dunn, R. L. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4639–4643.
3. DeJong, P. J., Grossovsky, A. J. & Glickman, B. W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3499–3503.
4. Barbacid, M. (1987) *Annu. Rev. Biochem.* **56**, 779–827.
5. Levine, A. J., Momand, J. & Finlay, C. A. (1991) *Nature (London)* **351**, 453–456.
6. Kunkel, T. A. (1985) *J. Biol. Chem.* **260**, 12866–12874.
7. Kunkel, T. A. & Alexander, P. S. (1986) *J. Biol. Chem.* **261**, 160–166.
8. Kunkel, T. A. & Bebenek, K. (1988) *Biochim. Biophys. Acta* **951**, 1–15.
9. Fry, M. & Loeb, L. A. (1986) *Animal Cell DNA Polymerases* (CRC, Boca Raton, FL).
10. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
11. Williams, K. J., Loeb, L. A. & Fry, M. (1990) *J. Biol. Chem.* **265**, 18682–18689.
12. Perrino, F. W. & Loeb, L. A. (1989) *J. Biol. Chem.* **264**, 2898–2905.
13. Mendelman, L. V., Petruska, J. & Goodman, M. F. (1990) *J. Biol. Chem.* **265**, 2338–2346.
14. Bebenek, K., Joyce, C. M., Fitzgerald, M. P. & Kunkel, T. A. (1990) *J. Biol. Chem.* **265**, 13878–13887.
15. Fay, P. J., Johnson, K. O., McHenry, C. S. & Bambara, R. A. (1982) *J. Biol. Chem.* **257**, 5692–5699.
16. Kaguni, L. S. & Clayton, D. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 983–987.
17. Weaver, D. T. & DePamphilis, M. L. (1982) *J. Biol. Chem.* **257**, 2075–2086.
18. Hillbrand, G. G. & Beattie, K. L. (1985) *J. Biol. Chem.* **260**, 3116–3125.
19. Abbotts, J., SenGupta, D. N., Zon, G. & Wilson, S. H. (1989) *J. Biol. Chem.* **263**, 15094–15103.
20. Weisman-Shomer, P., Dube, D. K., Perrino, F. W., Stokes, K., Loeb, L. A. & Fry, M. (1989) *Biochem. Biophys. Res. Commun.* **164**, 1149–1159.
21. Gross, F. & Krauss, G. (1984) *Eur. J. Biochem.* **141**, 109–114.
22. Fry, M., Weisman-Shomer, P., Lapidot, J. & Sharf, R. (1987) *J. Biol. Chem.* **262**, 8861–8867.
23. Sharf, R., Weisman-Shomer, P. & Fry, M. (1988) *Biochemistry* **27**, 2990–2997.
24. Asna, N., Weisman-Shomer, P. & Fry, M. (1989) *J. Biol. Chem.* **264**, 5245–5252.
25. LaDuca, R. J., Fay, P. J., Chuang, C., McHenry, C. S. & Bambara, R. A. (1983) *Biochemistry* **22**, 5177–5188.
26. Kaguni, L. S., DiFrancesco, R. A. & Lehman, I. R. (1984) *J. Biol. Chem.* **259**, 9314–9319.
27. Tabor, S., Huber, H. E. & Richardson, C. C. (1987) *J. Biol. Chem.* **262**, 16212–16223.
28. Huber, H. E., Tabor, S. & Richardson, C. C. (1987) *J. Biol. Chem.* **262**, 16224–16232.
29. Hopfield, J. J. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5248–5252.
30. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H. & Kunkel, T. A. (1989) *J. Biol. Chem.* **264**, 16948–16956.
31. Perrino, F. W., Preston, B. D., Sandell, L. L. & Loeb, L. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8343–8347.