Influence of Local Nucleotide Sequence on Substitution of 2-Aminopurine for Adenine during Deoxyribonucleic Acid Synthesis in Vitro†

Reynaldo C. Pless‡ and Maurice J. Bessman*

ABSTRACT: Three highly purified DNA polymerases, *Escherichia coli* polymerase I (enzyme A) and the polymerases induced by wild-type T4 phage and by T4 phage mutant L141 (antimutator phenotype), have been examined with respect to their tendency to incorporate the deoxyribonucleotide of 2-aminopurine [(AP)] for deoxyadenylate at specific sites in deoxyribonucleic acid (DNA). Using ϕX174 phage DNA as a template and selected ϕX174 restriction fragments as specific primers, we synthesized short sequences of ϕX174 DNA in vitro by the polymerase of interest, with the 5′-triphosphate of 2-aminopurine deoxyriboside and dATP at equimolar concentration. The relative incorporation of (AP) at the various adenine sites was determined by providing the newly synthesized DNA fragment with a specific terminal radioactive label, subjecting the DNA fragment to thermal depurination as a DNA cleavage reaction highly selective for (AP), and analyzing the resulting radioactive fragments by denaturing gel electrophoresis, autoradiography, and microdensitometry. The L141 polymerase shows very pronounced site-dependent variations in (AP) incorporation. For the wild-type T4 polymerase, the pattern of (AP) incorporation follows the biases seen for the L141 enzyme, although in a less pronounced form. Sequence preferences for (AP) incorporation are least marked for *E. coli* polymerase I (enzyme A); in several instances, they run counter to the sequence biases observed with the T4 enzymes. For the enzyme showing the most pronounced sequence effects, L141 polymerase, the extent of (AP) incorporation was determined at 57 different sites. No simple principle governing the sequence dependence of (AP) incorporation could be deduced from these results.

The widely divergent susceptibility of different loci of the rII region of the bacteriophage T4 to the mutagenic action of 2-aminopurine [(AP)][1] was documented in 1959 by Freese. Subsequent investigations (Koch, 1971; Salts & Ronen, 1971; Ronen et al., 1976, 1978; Coulondre et al., 1979) strongly suggested an influence of the local DNA sequence on the mutability of a given site by (AP). On the assumption that this sequence dependence of the mutagenic action of (AP) might, at least in part, be accounted for by different sequence-dependent propensities for misincorporation of d(AP)MP at A sites in newly synthesized DNA, we have recently determined the misincorporation of (AP) for Ade during DNA synthesis in vitro as a function of the preceding base, i.e., by nearest-neighbor analysis (Pless et al., 1981). DNA polymerases containing a highly active 3′-exonuclease, particularly the polymerase induced by the T4 antimutator phage L141, displayed a marked site selectivity in their incorporation of (AP), with a pronounced bias for incorporation of (AP) after Gua and against incorporation of (AP) after Ade. To examine the dependence of (AP) incorporation on the local base sequence in greater detail, we have now quantitated the enzyme-catalyzed incorporation of (AP) for Ade at specific sites in known DNA sequences. To this end, we have developed an (AP)-selective chemical cleavage procedure, analogous to the base-selective cleavages used in the Maxam–Gilbert sequencing technique (Maxam & Gilbert, 1977). In contrast to the qualitative nature of a DNA sequence determination, the quantitative character of our proposed analysis required that the following special conditions be met:

(1) The selectivity of the cleavage reaction for (AP) moieties in the DNA had to be very high, since the purpose of the investigation was to analyze a *chemically inhomogeneous* DNA, quantitating a low level of (AP) misincorporation for Ade [an average of 5% in the case of the highly selective L141 polymerase, at 30 °C and equimolar input of dATP and d(AP)TP; Pless et al., 1981]; this requires a highly selective cleavage reaction for the (AP) signal to be sufficiently prominent over the background of cleavage at other bases to allow a meaningful quantitation.

(2) The susceptibility of d(AP)MP residues and of dAMP residues in the DNA to the cleavage reaction should be independent of the local sequence, so that observed differences in the cleavage signals can be unambiguously ascribed to different extents of (AP) incorporation.

(3) When the cleavage reaction is performed conditions conducive to multiple cleavages in the same strand should be avoided, as they would distort the results in the direction of an artificially high incidence of (AP) in A positions proximal to the labeled end. Ideally, cleavage patterns ought to be analyzed in the limit of vanishing cleavage.

(4) The substitution of (AP) nucleotides for adenine nucleotides in a DNA fragment should ideally leave the mobility of the fragment in denaturing gel electrophoresis unaffected; otherwise, the electrophoretic separation might be impaired by considerable band broadening due to the fact that in our experiments the radioactively labeled isostichs are chemically inhomogeneous, containing adenines and 2-aminopurines distributed over the A sites in many different permutations.

The present report describes a procedure for (AP)-selective cleavage and its application to the quantitation of (AP) in-

† From the McCollum-Pratt Institute and Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218. *Received December 30, 1982; revised manuscript received June 21, 1983.* This is Contribution No. 1218 from the McCollum-Pratt Institute. This work was supported by Grant GM18649 from the National Institutes of Health.

‡ Present address: Department of Chemistry, The Ohio State University, Columbus, OH 43210.

[1] Abbreviations: (AP), 2-aminopurine; (AP)dR, 2-aminopurine deoxyriboside; d(AP)MP, d(AP)DP, and d(AP)TP, the 5′-monophosphate, 5′-diphosphate, and 5′-triphosphate of (AP)DR, respectively; dAdo, deoxyadenosine; dGuo, deoxyguanosine; dCyd, deoxycytidine; dThd, thymidine; Ade, adenine; Gua, guanine; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid.

0006-2960/83/0422-4905$01.50/0

corporation at defined A sites in known sequences of ϕX174 DNA, synthesized with a variety of polymerases, under conditions in which dATP and d(AP)TP compete for the same A sites.

Materials and Methods

Materials. Chemicals were obtained from the following sources: piperidine, Fisher Scientific; acrylamide and *N,N'*-methylenebis(acrylamide), Bethesda Research Laboratories, Inc.; urea, J. T. Baker Chemicals; [α-^{32}P]dATP (2000 Ci/mmol), New England Nuclear; nonradioactive nucleosides and nucleotides, P-L Biochemicals. (AP)dR, d(AP)TP, and [6-^{3}H]d(AP)TP were the materials described previously (Pless et al., 1981). [γ-^{32}P]ATP was prepared by the method of Glynn & Chappell (1964) as modified by Weiss et al. (1968). d-$(pT)_3$ and d-A$(pA)_3$ were kindly supplied by Dr. P. S. Miller.

Sources for the enzymes were the following: bacterial alkaline phosphatase, Worthington; *Micrococcus luteus* polymerase, Miles Laboratories, Inc.; T4 polynucleotide kinase and *E. coli* polymerase I (enzyme A), Boehringer Mannheim; restriction endonucleases *Hin*fI and *Msp*I, New England Biolabs, Inc. Terminal deoxynucleotidyltransferase was fraction 4' as described by Bollum et al. (1974). The purification of the wild-type T4 polymerase (T4D) and of the antimutator polymerase (L141) has been described (Lo & Bessman, 1976). In both cases, the homogeneous enzyme (fraction V) was used.

For all DNA polymerases employed, 1 unit incorporates 10 nmol of nucleotide into DNA at 37 °C in 30 min. One unit of terminal deoxynucleotidyltransferase incorporates 1 nmol of dATP into an acid-precipitable product with d-$(pT)_3$ as initiator, at 37 °C in 60 min (Bollum et al., 1974).

ϕX174 *am3* phage DNA and ϕX174 *am3* replicative form DNA were prepared as described by Cunningham et al. (1980). The requisite ϕX174 *am3* and *E. coli C* strain HF 4707 (Benbow et al., 1971) were obtained from Dr. R. P. Cunningham.

General Techniques. UV absorption spectra were measured with a Beckman DU monochromator fitted with a Gilford 252 photometer. Descending paper chromatography was carried out on Whatman 3MM sheets. Hydrolysis in 0.5 M aqueous piperidine at elevated temperatures was performed in sealed glass capillaries. Preparative gel electrophoresis of double-stranded DNA fragments as well as gel electrophoresis on thin denaturing gels was carried out according to Maxam & Gilbert (1980). DNA was extracted from polyacrylamide gel slices by electroelution as described by Danna (1980). Autoradiography was performed with Kodak X-omat R film. Film development was as given by Swanstrom & Shank (1978).

For the quantitative analysis of an autoradiogram, the developed film was scanned down the center of the electrophoretic lane with a Joyce-Loebl MK III C densitometer. The effective width of the scanning was 2.5 mm. As the electrophoretic channels were usually wider than that, parallel scans were executed, offset by 2.5 mm, and each radioactive band was quantitated by summing over the peak areas with which the band appeared in the various parallel scans. Peak areas were determined by paper cutting and weighing.

Preparation of d-$(pT)_3p(AP)$ and d-$(pT)_3[p(AP)]_2$. d-$(pT)_3$ (1.08 μmol of trinucleotide) and [6-^{3}H]d(AP)TP (1.62 μmol, 6.0×10^6 cpm/μmol) in 1.00 mL of 0.2 M potassium cacodylate, pH 7.2, 4 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 100 μg/mL bovine serum albumin were incubated with 2300 units of terminal deoxynucleotidyltransferase at 37 °C for 3 h. The incubation mixture was fractionated on a column of DEAE-Sephadex (HCO_3^- form) using a linear gradient of NH_4HCO_3, pH 8.0. The eluate was monitored by UV absorbance measurements and scintillation counting of aliquots. Appropriate fractions were pooled [d-$(pT)_3$p(AP), 79 nmol of oligomer, eluting at 0.20 M NH_4HCO_3; d-$(pT)_3[p(AP)]_2$, 34 nmol of oligomer, eluting at 0.24 M NH_4HCO_3] and evaporated to dryness.

Preparation of [5'-^{32}P]d-$(pT)_3p(AP)$, [5'-^{32}P]d-$(pT)_3[p(AP)]_2$, and [5'-^{32}P]d-$(pA)_4$. The 5'-terminal phosphate of d-$(pT)_3$p(AP) was cleaved with bacterial alkaline phosphatase. The product was purified on a DEAE-Sephadex column using a linear gradient of NH_4HCO_3, pH 8.0, and desalted by passage through a column of Sephadex G-15, in water. The material was converted to [5'-^{32}P]d-$(pT)_3$p(AP) by treatment with [γ-^{32}P]ATP and T4 polynucleotide kinase as described by Maxam & Gilbert (1980).

[5'-^{32}P]d-$(pT)_3[p(AP)]_2$ and [5'-^{32}P]d-$(pA)_4$ were prepared in analogous fashion from d-$(pT)_3[p(AP)]_2$ and d-A$(pA)_3$.

In Vitro Synthesis of Fragment ϕX174/HinfI-16b with Various Polymerases and Analysis of the Distribution of Incorporated (AP). Preparation and isolation of *Hin*fI restriction fragments from ϕX174 replicative form DNA were performed according to Maxam & Gilbert (1980).

(A) 3' End Labeling of the Priming Fragment. Twenty picomoles of restriction fragment *Hin*fI-1 in 10 μL of 67 mM Tris-HCl, pH 7.5, 3.3 mM $MgCl_2$, and 1 mM 2-mercaptoethanol was incubated with 44 pmol of [α-^{32}P]dATP and 3 units of *M. luteus* polymerase for 15 min at 15 °C. This resulted in the conversion of 45% of the radioactivity to an acid-insoluble form. The desired radioactively labeled *Hin*fI-1 fragment was purified by preparative electrophoresis through a polyacrylamide gel, electroelution, phenol extraction, and ethanol precipitation.

(B) Annealing of the Priming Fragment to ϕX174 Phage DNA. The pellet (approximately 20 pmol of 3' end-labeled fragment *Hin*fI-1) was taken up in 25 μL of 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA containing 20 pmol of ϕX174 phage DNA. To effect denaturation, 3 μL of 0.75 M NaOH was added, and the solution was incubated at 37 °C for 20 min. To allow reannealing, the pH was adjusted to 8.1 by addition of 4 μL of 0.75 M Tris–0.96 M HCl, and the solution was incubated for 30 min at 60 °C and for 30 min at 50 °C and chilled at 0 °C. Analysis of an aliquot by gel electrophoresis and autoradiography indicated approximately one-fourth of the total radioactivity to be associated with the viral DNA (remaining at the origin), as expected, indicating formation of approximately 10 pmol of the desired template-primer.

(C) Incubation with DNA Polymerase. The solution was adjusted to contain 0.020 μM template-primer in 67 mM Tris, pH 8.8, 6.7 mM $MgCl_2$, 17 mM $(NH_4)_2SO_4$, 2 mM NaCl (arising from the annealing process), 10 mM 2-mercaptoethanol, 160 μg/mL bovine serum albumin, 50 μM each of dCTP, dGTP, dTTP, and dATP, and 40 μM (in the incubation with the *E. coli* polymerase) or 60 μM (in the case of the T4 phage polymerases) d(AP)TP. Aliquots (100 μL) of this solution (containing 2 pmol of template-primer) were separately incubated with 1 unit of each of the polymerases of interest: *E. coli* polymerase I (enzyme A), L141 (antimutator) polymerase, and wild-type T4D polymerase, at 15 °C for 7 h.

(D) Isolation of the Newly Synthesized Fragment HinfI-16b. After extraction with phenol and ether, and ethanol precipitation, the polymerase reaction mixture was digested with 5 units of *Hin*fI restriction endonuclease in 50 μL of 50 mM NaCl, 6 mM Tris-HCl, pH 7.4, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol, and 100 μg/mL bovine serum albumin,

for 3 h at 37 °C. After extraction with phenol and ether, and ethanol precipitation, the DNA was fractionated by electrophoresis on an 11-cm gel. Autoradiography revealed a strong band at the position corresponding to the priming fragment *Hin*fI-1, a somewhat weaker band at the origin (presumably radioactive species hybridized to full-length ϕX174 DNA), and a small amount of radioactive material at the position corresponding to the desired 40-mer, *Hin*fI-16b. The latter band was electroeluted and precipitated with ethanol, and the dried residue was dissolved in 10 μL of water.

(*E*) (*AP*)-*Selective Hydrolysis*. The 10-μL aqueous sample of 5′ end-labeled *Hin*fI-16b was diluted with 10 μL of 1 M aqueous piperidine and heated to 90 °C for 18 h. The solution was diluted with 20 μL of water and lyophilized to dryness. After two additional lyophilizations from 20 μL of water, the residue was loaded onto a sequencing gel and electrophoresed as described by Maxam & Gilbert (1980). This was followed by autoradiography and densitometric quantitation of the film.

Results

Initial studies designed to identify a reaction with an ultimate potential for (AP)-selective DNA cleavage were carried out at the level of the deoxyribonucleosides. Comparing the stability of (AP)dR, dAdo, and dGuo to hydrolysis at 90 °C, we observed that in the entire pH range examined by us, from pH 7 to pH 13, (AP)dR is distinctly more labile than dAdo or dGuo. The reactions are conveniently monitored by measuring the UV absorption spectra in aqueous alkaline solution (pH 12). In this medium, it became evident that the UV-absorbing product of the hydrolysis of (AP)dR at 90 °C is the free base: the ratio A_{300}/A_{280}, which is 2.36 for the nucleoside, moved far toward the value of 1.34 which is characteristic for free 2-aminopurine.

That the primary result of hydrolysis of (AP)dR at elevated temperatures is cleavage of the glycosidic linkage was also evident from paper chromatography in 10 mM glycine, pH 10. From visual examination of the chromatograms under UV light, we estimated the rate of glycosidic cleavage of (AP)dR to be approximately 10-fold higher than the rate of cleavage of dAdo and dGuo, and higher yet compared to the rate of cleavage of dCyd and dThd. Thus, these exploratory experiments augured well for the use of hydrolysis at elevated temperatures for (AP)-selective cleavage, as it affected (AP)dR with high selectivity and as it led to depurination, which is the usual prelude to cleavage of the DNA backbone.

As the standard method for hydrolysis, we chose 0.5 M aqueous piperidine at 90 °C, since these were the conditions used by Maxam and Gilbert in 1977 for effecting the β-eliminations leading to the break in the polynucleotide backbone *after* the initial base-selective labilization or rupture of the glycosidic linkages, implying that the unaffected nucleotide residues remain sufficiently resistant to this treatment. Also, when applied to polynucleotides, the hot aqueous piperidine treatment would bring about the *entire* sequence of reactions leading to DNA cleavage at (AP) sites, viz., depurination *plus* piperidine-catalyzed β-eliminations.

To examine the selective cleavage at (AP) residues at the polynucleotide level, we prepared the model oligomers [5′-^{32}P]d-(pT)$_3$p(AP), [5′-^{32}P]d-(pT)$_3$[p(AP)]$_2$, and, for comparison, [5′-^{32}P]d-(pA)$_4$ as described under Materials and Methods. These compounds were hydrolyzed in 0.5 M aqueous piperidine at 90 °C for varying lengths of time, and the hydrolysates were analyzed by electrophoresis through a 25% polyacrylamide gel, followed by autoradiography. Figure 1 presents the densitometric tracings. The 24-h hydrolysate of [5′-^{32}P]d-(pT)$_3$p(AP) contains a major, faster moving

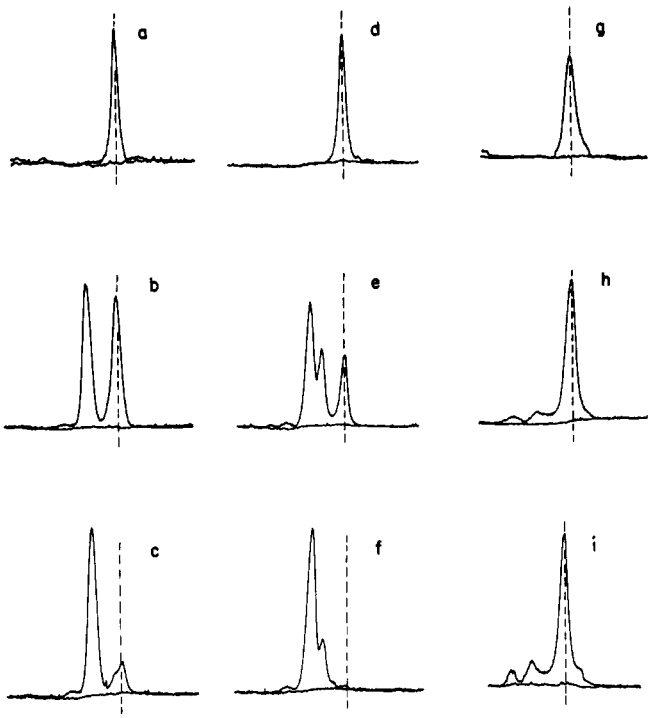


FIGURE 1: Densitometric tracings of autoradiographed gel electrophoretograms of 5′-^{32}P-labeled model oligonucleotides, hydrolyzed in 0.5 M aqueous piperidine at 90 °C for different times: d-p*TpTpTp(AP) hydrolyzed for (a) 0, (b) 24, or (c) 72 h; d-p*TpTpTp(AP)p(AP) hydrolyzed for (d) 0, (e) 24, or (f) 72 h; d-p*ApApApA hydrolyzed for (g) 0, (h) 24, or (i) 72 h. The direction of electrophoresis was from right to left. The dashed lines indicate the electrophoretic position of the starting material.

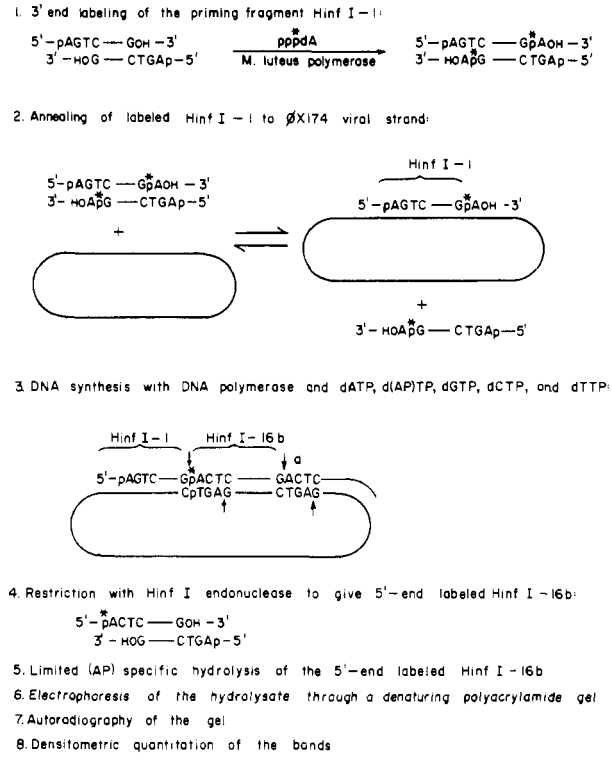
product, identified as [5′-^{32}P]d-(pT)$_3$p on the basis of its comigration with material obtained from [5′-^{32}P]d-(pT)$_3$p(AP) by depurination with diphenylamine–formic acid (Burton & Petersen, 1960). This is the expected product, arising from depurination followed by removal of the 3′-terminal sugar. After a 72-h hydrolysis, the [5′-^{32}P]d-(pT)$_3$p band is clearly the most prominent, and the starting material, [5′-^{32}P]d-(pT)$_3$p(AP), is almost depleted.

The 24-h hydrolysis of [5′-^{32}P]d-(pT)$_3$[p(AP)]$_2$ again produces [5′-^{32}P]d-(pT)$_3$p as the fastest moving product, arising from cleavage at the internal (AP) residue, along with a slower moving product, interpreted as [5′-^{32}P]d-(pT)$_3$p(AP)p, from cleavage at the terminal (AP) position. After 72 h, the [5′-^{32}P]d-(pT)$_3$p band has further increased, partly at the expense of [5′-^{32}P]d-(pT)$_3$p(AP)p; the starting material, [5′-^{32}P]d-(pT)$_3$[p(AP)]$_2$, has been entirely consumed. This tracing also shows the presence of a small amount of a species moving faster than d-(pT)$_3$p, probably [5′-^{32}P]d-(pT)$_2$p, indicating a low extent of cleavage at T sites under the conditions of the hydrolysis.

In the case of [5′-^{32}P]d-(pA)$_4$, the starting material clearly remained the predominant constituent of the hydrolysate for at least 72 h. Small amounts of faster moving species, interpreted as [5′-^{32}P]d-(pA)$_3$p and [5′-^{32}P]d-(pA)$_2$p, are apparent after a 24-h hydrolysis; these bands become more pronounced after 72 h. Thus, in hot aqueous piperidine, oligodeoxyribonucleotides are clearly more susceptible to cleavage at Ade residues than at Thy; however, Ade positions are distinctly more resistant to hydrolysis than are (AP) positions.

From the densitometric tracings for the 24-h hydrolysates, we estimate first-order rate constants of 125×10^{-7}, 160×10^{-7}, and 36×10^{-7} s^{-1} for the disappearance of d-(pT)$_3$p(AP),

Scheme I: Synthesis of ϕX174 *Hin*fI-16b with dATP and d(AP)TP in Competition and Analysis of the Distribution of (AP)



[a] Arrows indicate *Hin*fI cleavage sites.

d-$(pT)_3[p(AP)]_2$, and d-$(pA)_4$, respectively, corresponding to 125×10^{-7} s^{-1} per (AP) residue, 80×10^{-7} s^{-1} per (AP) residue, and 9×10^{-7} s^{-1} per Ade residue, respectively. These rate constants are similar to those estimated by us for the hydrolysis of the free deoxyribonucleosides, (AP)dR and dAdo, in 0.5 M piperidine at 90 °C.

These experiments using model oligonucleotides indicated that hydrolysis in 0.5 M aqueous piperidine at 90 °C was a suitable procedure for effecting DNA cleavage selectively at (AP) positions. This method was then used to quantitate the (AP) incorporated at various A sites in known DNA sequences synthesized by different polymerases under conditions in which Ade and (AP) competed for the same sites. As template, we selected the fully sequenced, single-stranded DNA of ϕX174 (Sanger et al., 1978). Template-dependent DNA polymerization would be primed by the complementary strand of a selected ϕX174 restriction fragment. The experimental protocol is outlined in Scheme I. The ϕX174 restriction fragment, e.g., *Hin*fI-1, 726 base pairs long, is incubated with *M. luteus* polymerase and [α-^{32}P]dATP. This results in extension of the 3′ termini by one p*dA unit, both in the viral (plus) strand and in the complementary (minus) strand. The labeled fragment is isolated by gel electrophoresis and incubated with ϕX174 viral form DNA under denaturing conditions. Upon renaturation, the restriction fragment will in part reanneal in its original form; no extended synthesis is possible on this duplex. In part, the minus strand of the labeled fragment will anneal to its complementary region on the circular viral DNA; the resulting primed DNA is capable of sustaining extensive DNA synthesis in the desired direction. The normal complement of triphosphates and d(AP)TP to compete with dATP are now added, and the mixture is incubated with the polymerase of interest. To the extent that DNA synthesis reconstitutes the *Hin*fI recognition sequence at the initial priming site in its double-stranded form, the product will become cleavable again by *Hin*fI endonuclease. The radioactive tag, which resided in the 3′-terminal nucleotide of the primer, will, upon this cleavage, emerge as the 5′-terminal phosphate of the newly synthesized complementary strand DNA, a suitable end label for sequencing purposes. To the extent that synthesis carries over the next *Hin*fI recognition site, which in this example occurs 40 nucleotides downstream from the first cleavage site, the DNA will become cleavable at this site as well. After polymerization and restriction, the desired product, which in our example is the 5′-^{32}P-labeled 40-mer *Hin*fI-16b, is isolated by gel electrophoresis and then subjected to the (AP)-selective cleavage reaction with hot aqueous piperidine. The hydrolysate is resolved on a sequencing-type polyacrylamide gel (Maxam & Gilbert, 1980). The gel is autoradiographed, and the autoradiogram is quantitatively evaluated by densitometry. The assignment of peaks was based on the known sequence of ϕX174 DNA, the migration of the bands relative to the dyes xylene cyanol and bromophenol blue run in the same channel, and the parallel electrophoresis of (G + A)-specific cleavage reaction mixtures (Maxam & Gilbert, 1980) of the same fragment.

Three highly purified DNA polymerases were examined in this manner: the DNA polymerases induced by wild-type T4 phage and by the T4 antimutator phage L141 as well as the proteolytic fragment of *E. coli* polymerase I which contains the polymerase and 3′–5′-exonuclease activities ["enzyme A" in the nomenclature of Klenow et al. (1971); "large fragment" according to Setlow et al. (1972)]. Experiments using intact *E. coli* polymerase I resulted in complete loss of radioactive label, probably due to the 5′–3′-exonuclease activity of the enzyme digesting the entire priming fragment from its 5′ terminus. Loss of label by the action of the 3′–5′-exonuclease on the 3′-terminal nucleotide at the start of the polymerization reaction was not monitored; it can be assumed to be substantial in the case of the L141 polymerase (Muzyczka et al., 1972; Bessman et al., 1974). The temperature chosen for DNA synthesis was 15 °C, in analogy to similar experiments by Weymouth & Loeb (1978) on the action of *E. coli* polymerase I in vitro. Controls showed that under these conditions net incorporation of [^{3}H]dTMP was still increasing at reaction times much longer than the 7-h reaction time used in our experiment.

Figure 2 shows the autoradiographed gel electrophoretograms for partial hydrolysates of the 5′ end-labeled 40-mer *Hin*fI-16b, synthesized under a variety of conditions. In each case, hydrolysis was in 0.5 M aqueous piperidine, at 90 °C, for 18 h. In this electrophoresis, the shortest oligomers (up to 5′-p*ACTp) were eluted off the gel; the shortest oligonucleotide remaining on the gel is 5′-p*ACTCp, corresponding to the first A position for which Ade and (AP) competed during DNA synthesis. In lanes 1, 2, and 3, the 40-mer was synthesized by the action of *E. coli* polymerase I (enzyme A), antimutator L141 polymerase, and wild-type T4 polymerase, respectively, at 15 °C, under conditions where Ade and (AP) were competing for the same sites. The input concentrations of d(AP)TP were 80% of the input concentration of dATP in the case of *E. coli* polymerase I (enzyme A) and 120% of the input concentration of dATP with the T4 enzymes. The strongest bands in the electrophoretograms correspond to the A positions, as expected. In the case of *E. coli* polymerase I (enzyme A) (lane 1), the intensities of the A bands are approximately equal, indicating a similar extent of substitution of (AP) for Ade at these positions. In contrast, for the 40-mer synthesized by the antimutator L141 polymerase (lane 2), the A bands display widely different intensities; we interpret the very weak bands as denoting positions of particularly low

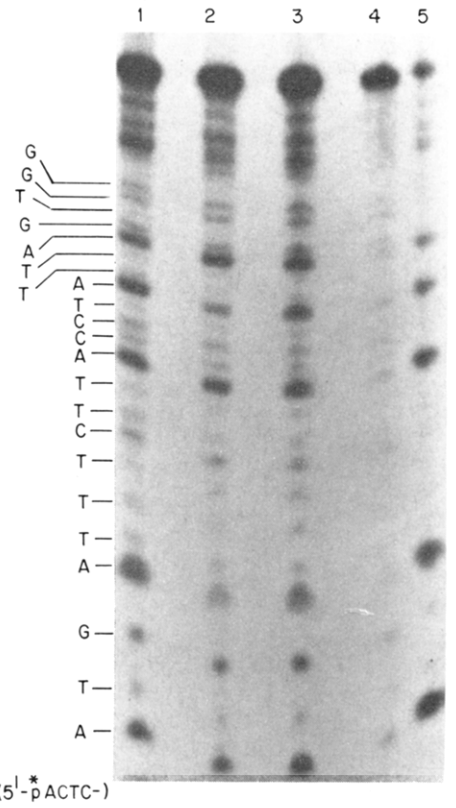


FIGURE 2: Autoradiographed gel electrophoretograms of 5′-^{32}P-labeled *Hin*fI-16b fragments (complementary strand) after hydrolysis in 0.5 M aqueous piperidine at 90 °C for 18 h. The *Hin*fI-16b fragments were synthesized with dGTP, dCTP, dTTP, and, in addition, (lane 1) dATP and d(AP)TP with *E. coli* polymerase I (enzyme A), (lane 2) dATP and d(AP)TP with L141 (antimutator) polymerase, (lane 3) dATP and d(AP)TP with wild-type T4 polymerase, (lane 4) dATP with *E. coli* polymerase I (enzyme A), and (lane 5) d(AP)TP with *E. coli* polymerase I (enzyme A).

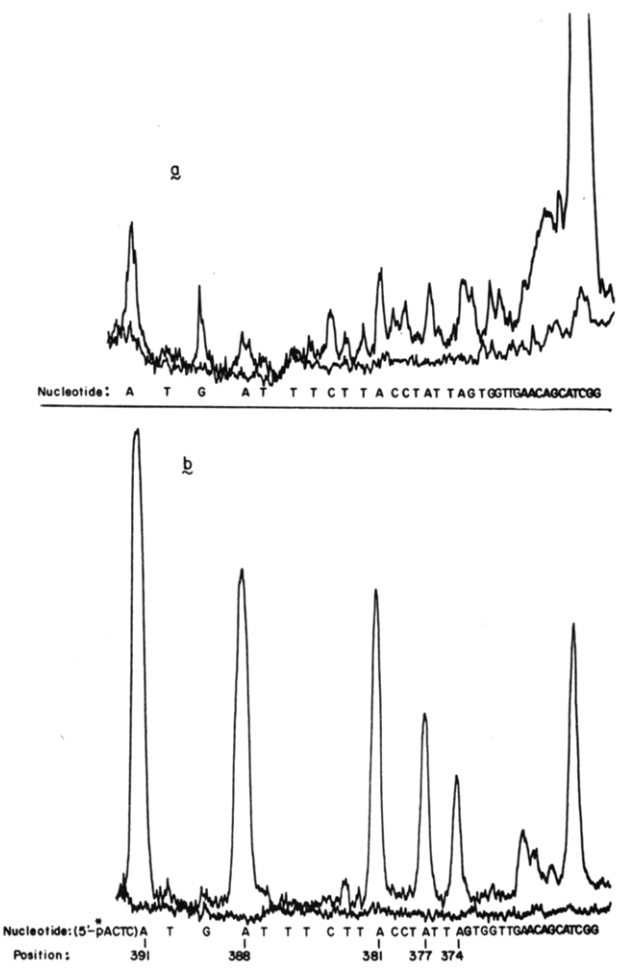


FIGURE 3: Densitometric scans of autoradiographed gel electrophoretograms of 5′-^{32}P-labeled *Hin*fI-16b fragments after hydrolysis in 0.5 M aqueous piperidine at 90 °C for 18 h. Nucleotide position is as numbered in the sequence of Sanger et al. (1978). (a) *Hin*fI-16b in which all A sites are occupied by Ade; (b) *Hin*fI-16b in which all A sites are occupied by (AP). The direction of electrophoresis is from right to left. The peak on the extreme right is the parent 40-mer *Hin*fI-16b.

substitution of (AP) for Ade. The hydrolysate of the 40-mer synthesized with wild-type T4 polymerase (lane 3) shows the same pattern of weak and strong A bands as observed with the L141 polymerase, although in a less pronounced form.

The autoradiogram also shows the hydrolysis patterns for two control *Hin*fI-16b 40-mers, synthesized by *E. coli* polymerase I (enzyme A): the first was prepared with a normal complement of deoxynucleoside triphosphates [without d(AP)TP] so that all A positions are occupied exclusively by Ade; the second 40-mer was obtained with d(AP)TP (at 350 μM concentration) fully replacing dATP in the incubation so that all A positions are exclusively occupied by (AP) [the fact that the 40-mer could be excised in this instance demonstrates that the sequence at the 3′-terminal end of the newly synthesized 40-mer—G(AP)CTC, with the opposite strand reading GAGTC—is recognized and cleaved by *Hin*fI endonuclease]. In the case of the normal complement DNA (lane 4), the parent peak, i.e., undegraded 40-mer, is by far the most plentiful constituent of the hydrolysate; only small amounts of degradation products, i.e., faster moving fragments, are present after an 18-h hydrolysis. In contrast, the 40-mer in which all A sites are occupied by (AP) (lane 5) is extensively degraded in an 18-h hydrolysis: the parent 40-mer is almost entirely consumed, and there is a pronounced increase in the intensity of the A bands as one goes to progressively shorter oligomers. This is a manifestation of multiple chain scission: obviously, in the limit of complete cleavage at all (AP) sites, the only radioactive oligonucleotide left would be 5′-p*ACTCp, which signals the position of the first (AP) proximal to the labeled end.

When due consideration is given to the convex curvature of the electrophoretic front, the oligonucleotides in the all-(AP) hydrolysate are seen to comigrate with the corresponding oligomers in the all-Ade hydrolysate, indicating that the substitution of (AP) for Ade does not appreciably alter the electrophoretic mobility of the fragments. This result is of importance because any significant fluctuation in the electrophoretic mobilities due to the chemical inhomogeneity of the isostichs would have seriously compromised the electrophoretic resolution upon which the analysis depends.

For a quantitative evaluation, the various channels were scanned with a densitometer. Figure 3a shows the profile obtained when the electrophoretogram corresponding to the all-Ade 40-mer is scanned down the center of the channel. The parent 40-mer is preeminent. The region containing the 29-mer and higher oligomers consists of an envelope of peaks; the shorter oligomers, however, are reasonably well resolved. Inspection of this region indicates that the DNA is most susceptible to cleavage at A sites and G sites. Quantitative evaluation of densitometric scans (as described under Materials and Methods) performed on this hydrolysate as well as on hydrolysates of other all-Ade fragments (data not shown) indicates that in partial hydrolysis in 0.5 M aqueous piperidine at 90 °C A sites and G sites are equally susceptible to cleavage. The propensity to cleavage of the various residues is inde-

pendent of sequence; i.e., in a given hydrolysate, all bands corresponding to a given base (e.g., all G bands) have approximately the same peak area.

The central scan for the 18-h hydrolysate of the all-(AP) 40-mer (Figure 3b) shows a drastically reduced parent 40-mer peak. The pattern of hydrolysis fragments is clearly dominated by the (AP) bands. A comparison of the intensities of these (AP) bands shows skewing toward the shortest oligomer. The first five (AP) bands, corresponding to positions 391, 388, 381, 377, and 374 in the ϕX174 sequence of Sanger et al. (1978), had peak areas of 103, 74, 44, 28, and 18 mg of paper, respectively. This skewing is clearly due to repeated scission within the same chain, a phenomenon which is certainly pervasive at a stage where the parent 40-mer is nearly consumed. Significantly, shorter hydrolyses of the all-(AP) 40-mer resulted in a more uniform weight of (AP) peaks: in a 5-h hydrolysate, the areas of the aforementioned five peaks were recorded as 54, 57, 50, 40, and 43 mg of paper, respectively, and a 1-h hydrolysate yielded 15, 17, 18, 18, and 16 mg. Thus, to a first approximation, the propensity to DNA cleavage at (AP) is independent of sequence.

From densitometric analysis of the electrophoretogram shown in Figure 3b, we obtain the following estimate of the susceptibility of DNA to cleavage at (AP) relative to susceptibility to cleavage at G: (AP):G = 15:1.

Figure 4 represents densitometric tracings corresponding to lanes 1–3 in Figure 2 [i.e., three different polymerases, with Ade and (AP) competing]. Five peaks (positions 391, 388, 381, 377, and 374) are rather well separated. In the case of *E. coli* polymerase I (enzyme A), the first four peaks appear in approximately equal intensity, while the fifth peak (A-374) is somewhat smaller. In comparison, the hydrolytic degradation pattern for the 40-mer synthesized by the antimutator L141 polymerase shows a large reduction in the intensity of the bands corresponding to A-388 and A-377. These two bands are also reduced, though not as markedly, in the case of the 40-mer synthesized by the wild-type T4 polymerase. The bands were quantitated as described under Materials and Methods. Peak areas (expressed in milligrams of paper weight) are listed in Table I for those peaks which are reasonably well resolved in the hydrolysates of the 40-mer synthesized by the three different enzymes. The table also lists, under the heading of (AP)/[A + (AP)], the percentage of the bases at the given A position which are made up by (AP), as calculated by interpolating the measured peak area between the values expected if the A position had been occupied exclusively by Ade (i.e., a peak area equal to the area observed at G sites) and the value expected if the A position had been occupied exclusively by (AP) (i.e., a peak area 15 times as large as the area observed at G sites), according to the formula

$$\frac{(AP)}{A + (AP)} = \frac{\text{area}[\text{Ade} + (AP)] - \text{area}(G)_{av}}{[15 \times \text{area}(G)_{av}] - \text{area}(G)_{av}} \times 100\%$$

To obtain a value for $(G)_{av}$, we averaged the areas of the well-resolved G peaks at positions 389, 371, and 370.

For the DNA synthesized by *E. coli* polymerase I (enzyme A), we note the approximate equivalence of the (AP) content of the A positions 391, 388, 381, and 377. The values, from 14% to 19%, indicate that at a slight disadvantage in substrate concentration [d(AP)TP/dATP = 0.8], (AP) is stably incorporated into these positions to about one-fifth of the extent of Ade incorporation. Position 374 shows a somewhat lower (AP) content, 9%. The incorporation of (AP) for Ade by the antimutator polymerase, L141, shows much wider differences from one A site to another, from little or no incorporation at

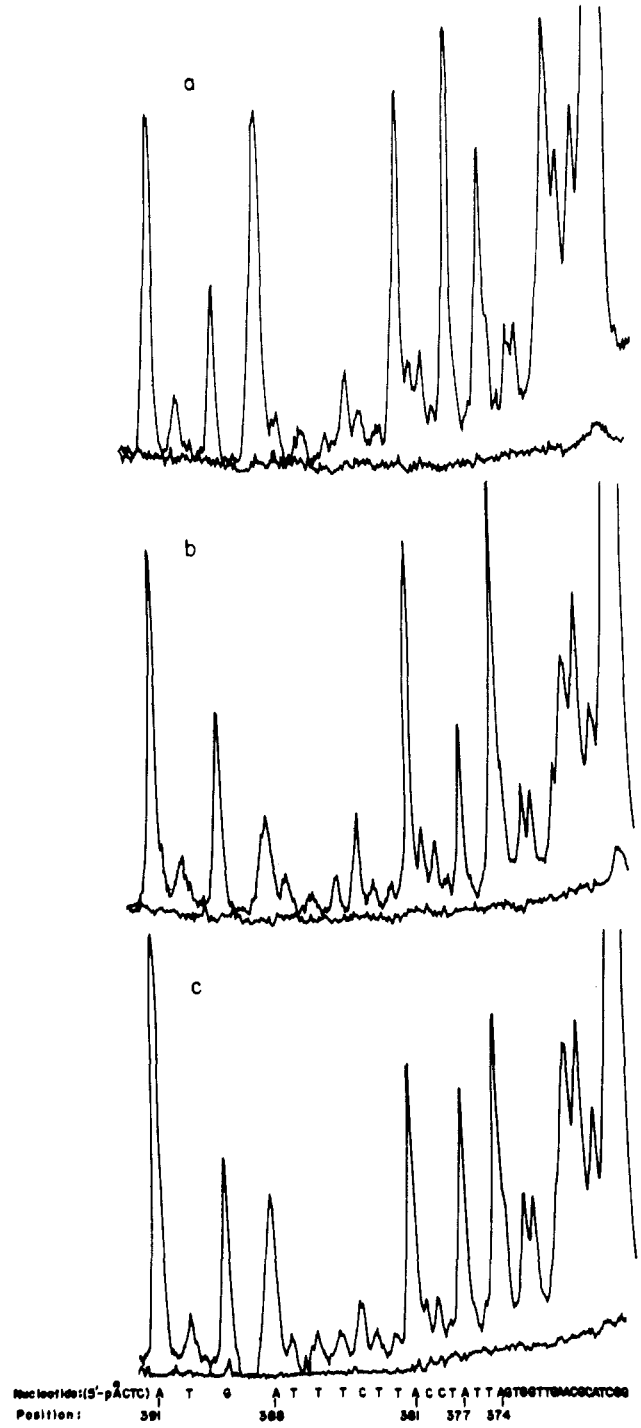


FIGURE 4: Densitometric scans of autoradiographed gel electrophoretograms of 5′-^{32}P-labeled *Hin*fI-16b fragments after hydrolysis in 0.5 M aqueous piperidine at 90 °C for 18 h. Nucleotide position is as numbered in the sequence of Sanger et al. (1978). The fragments had been synthesized by using the following enzymes [and the following input ratios d(AP)TP/dATP]: (a) *E. coli* polymerase I (enzyme A) (0.8); (b) L141 (antimutator) polymerase (1.2); and (c) wild-type T4 polymerase (1.2). The direction of electrophoresis was from right to left. The large peak on the right is the undegraded 40-mer *Hin*fI-16b.

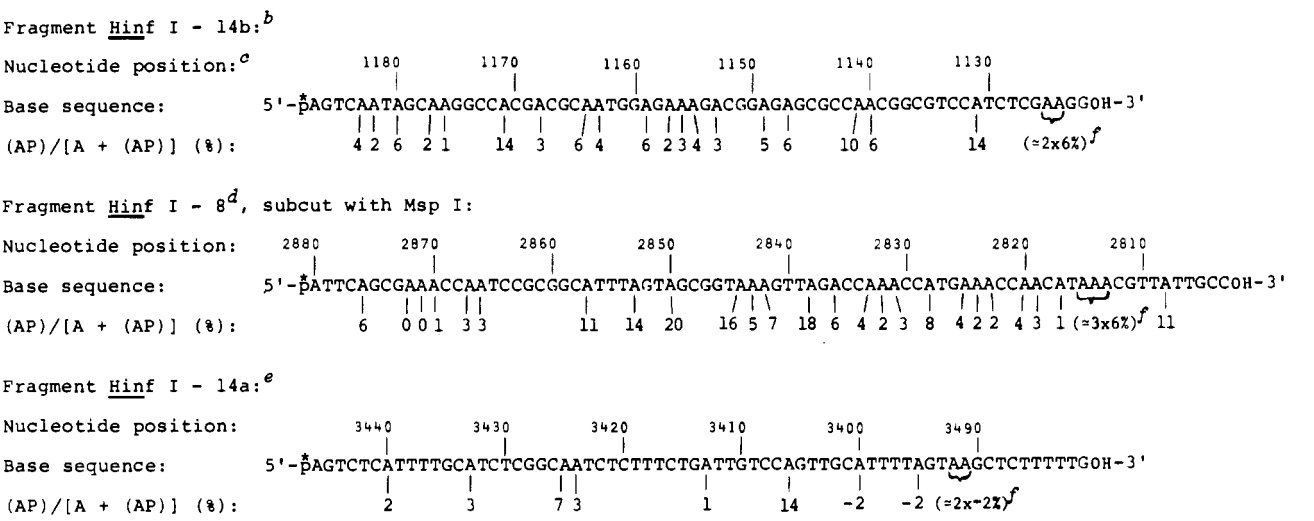
positions 388 (2%) and 377 (2%) to a substantial relative incorporation at 14% at position 374. The average relative incorporation of (AP) at these five sites with the L141 enzyme (8%) is only half as large as the average relative incorporation of (AP) with *E. coli* polymerase I (enzyme A) (15%), although the input concentration of d(AP)TP relative to dATP was higher in the synthesis with the L141 polymerase [d(AP)-TP/dATP = 1.2] than in the synthesis with polymerase I

Table I: Distribution of (AP) in *Hin*fI-16b Fragments Synthesized by Various DNA Polymerases

polymerase		*E. coli* Pol I (enzyme A)		L141		T4D		*E. coli* Pol I (enzyme A)	L141
input ratio, d(AP)TP/dATP		0.8		1.2		1.2		1.2	1.8
nucleotide position[a]	base sequence	peak area (mg)	(AP)/[A + (AP)] [b] (%)	peak area (mg)	(AP)/[A + (AP)] [c] (%)	peak area (mg)	(AP)/[A + (AP)] [d] (%)	(AP)/[A + (AP)] (%)	(AP)/[A + (AP)] (%)
(395–392)	(5′-p*ACTC-)[e]								
391	A	58	14	43	9	67	11	25	14
390	T	10		11		11			
389	G	20		23		30			
388	A	73	19	24	2	47	5	28	3
387	T	7		6		5			
386	T	10		6		8			
385	T	8		8		8			
384	C	15		15		14			
383	T	11		7		7			
382	T	9		6		7			
381	A	62	15	48	11	67	11	23	16
380	C	17		12		12			
379	C	18		13		12			
378	T	8		6		7			
377	A	67	17	23	2	56	8	28	5
376	T	9		5		7			
375	T	9		5		6			
374	A	(45)[g]	9	(56)[g]	14	(67)[g]	11	15	20
373	G	(20)[g]		(19)[g]		(27)[g]			
372	T	7		5		5			
371	G	20		18		27			
370	G	21		17		25			
(369–356)	(-TTGAACAGCA-)(-TCGGOH-3′)[f]								

[a] Nucleotide position as numbered in the ϕX174 DNA sequence of Sanger et al. (1978). [b] Calculated as [(peak area − 20 mg) × 100%]/[(15 × 20 mg) − 20 mg]. [c] Calculated as [(peak area − 19 mg) × 100%]/[(15 × 19 mg) − 19 mg]. [d] Calculated as [(peak area − 27 mg) × 100%]/[(15 × 27 mg) − 27 mg]. [e] The 5′-terminal sequence of the 40-mer; the oligonucleotides corresponding to the sequence were run off the gel. [f] The 3′-terminal sequence of the 40-mer; the oligonucleotides corresponding to the sequence were too ill-resolved to allow meaningful quantitation. [g] The areas for the A peak in position 374 were calculated by subtracting from the measured area of the unresolved band containing A-374 and G-373 the average area of G peaks in the same channel.

Table II: Distribution of (AP) in Three DNA Sequences Synthesized by L141 Polymerase[a]



[a] Synthesized with d(AP)TP/dATP = 1.2. [b] Primed by *Hin*fI-2. [c] Nucleotide position as numbered in the φX174 sequence of Sanger et al. (1978). [d] Primed by *Hin*fI-4. [e] Primed by *Hin*fI-12. [f] Less reliable values, due to poor electrophoretic resolution.

(enzyme A) [d(AP)TP/dATP = 0.8]. Another observation is that the distribution of sites of high and low incorporation of (AP) in the DNAs synthesized by these two enzymes do not coincide: of the five A sites under discussion, positions 388 and 377 show very low incorporation of (AP) with the L141 enzyme but higher than average incorporation of (AP) in the DNA synthesized with *E. coli* polymerase I (enzyme A); conversely, position 374, which shows a lower than average incorporation of (AP) in the case of the *E. coli* enzyme, is the site of the highest incorporation of (AP) with L141 polymerase.

Averaging over these five sites, we find that the relative incorporation of (AP) by the wild-type T4 polymerase is 9%. Here again, as in the case of L141 polymerase, positions 388 and 377 show the lowest relative incorporation of (AP), 5% and 8%, respectively, but the fluctuation in (AP) incorporation at the various A sites is less pronounced than in the DNA synthesized by the L141 enzyme.

Table I also presents the values for the relative incorporation of (AP) observed in a separate experiment, in which the same 40-mer was synthesized with *E. coli* polymerase I (enzyme A), with d(AP)TP and dATP at an input ratio of 1.2. The relative incorporation of (AP), averaged over the five sites, is 23%, or 1.6 times the average value of 14% observed with the same enzyme at an input ratio of 0.8. Comparing the values of (AP)/[A + (AP)] observed in the two experiments for corresponding peaks, one obtains factors of 1.8, 1.6, 1.5, 1.6, and 1.6 for positions 391, 388, 381, 377, and 374, respectively. Also shown in Table I are the values of (AP)/[A + (AP)] obtained when the experiment with L141 polymerase was repeated with the input ratio increased by a factor of 1.5 [d(AP)TP/dATP = 1.8]; this resulted in an increase in (AP) incorporation of 1.6, 1.5, 1.5, 2.5, and 1.4 at the five A positions.

With *E. coli* polymerase I (enzyme A) and with wild-type T4D polymerase, a highly increased relative incorporation of (AP) at A sites could be achieved by a large increase in the input ratio d(AP)TP/dATP. Figure 5 shows central densitometric scans documenting the 5′-terminal region of *Hin*fI-6 fragments, 311 base pairs long, the synthesis of which was primed by *Hin*fI-9 fragments. One of these fragments was synthesized by *E. coli* polymerase I (enzyme A), with d(AP)TP/dATP = 2.6 at input; the other fragment was synthesized by wild-type T4 polymerase, with d(AP)TP/dATP = 5.7. When compared to Figure 4, the scans in Figure 5 show a marked increase in the size of the peaks corresponding to A positions (relative to G peaks), illustrating that an increased incorporation of (AP) correlates with an increase in the input concentration of d(AP)TP in the polymerization mixture.

Attempts to examine DNA synthesis with the antimutator L141 polymerase at high concentrations of d(AP)TP were unsuccessful. Thus, polymerase incubations in which dATP and d(AP)TP were present in concentrations of 50 and 250 μM, respectively, failed to yield detectable amounts of the target fragments *Hin*fI-14b (a 66-mer) and *Hin*fI-8, subcut with *Msp*I (a 78-mer). It appears that, with L141 polymerase, a high ratio of d(AP)TP to dATP brings about a high level of aborted synthesis, so that the newly synthesized strand is never extended sufficiently to reach the next intended restriction site downstream. The aforementioned 66-mer and 78-mer, as well as fragment *Hin*fI-14a (a 66-mer), were successfully synthesized with the L141 enzyme at an input ratio of d(AP)TP/dATP = 1.2. Table II presents the distribution of (AP) in these sequences. Again, we observe a wide fluctuation in the extent to which the L141 polymerase incorporates (AP) at different A sites. In the case of *Hin*fI-14b, (AP)/[A + (AP)] varies between 1% and 14%, in the 78-mer between 0% and 20%, and in the case of *Hin*fI-14a between −2% and 14%.

Discussion

The purpose of the present work was to investigate the influence of local DNA sequence on the misincorporation of nucleotides by various DNA polymerases. The approach chosen was to use 2-aminopurine deoxyribonucleotide as a paradigm for the misincorporating nucleotide and to determine quantitatively the relative incorporation of (AP) at the different A sites in defined DNA sequences synthesized by the polymerase of interest. The method described in this report, based on end labeling of the newly synthesized strand, selective DNA cleavage at (AP) moieties, and analysis of the resulting radioactive fragments, provides for the required quantitative determination of the distribution of (AP) in the DNA.

The reliability of the method can be gauged by comparing the values of (AP)/[A + (AP)] at given A sites obtained in separate experiments with the same polymerase, given in Table I. The results of the repeat experiments (last two columns)

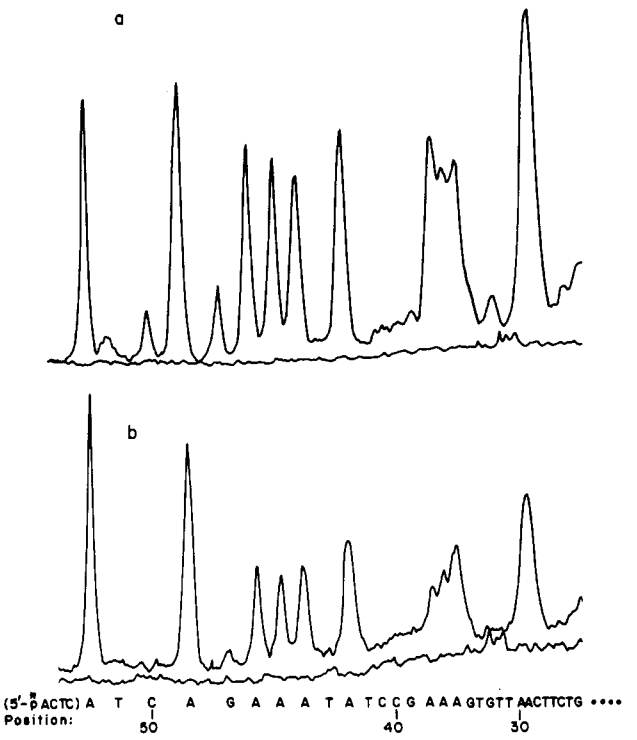


FIGURE 5: Partial densitometric scans of autoradiographed gel electrophoretograms of 5′-^{32}P-labeled *Hin*fI-6 fragments after hydrolysis in 0.5 M aqueous piperidine at 90 °C for 5 h. Nucleotide position is as numbered in the ϕX174 sequence of Sanger et al. (1978). The fragments had been synthesized with (a) *E. coli* polymerase I (enzyme A) in the presence of dATP (50 μM) and d(AP)TP (130 μM) and (b) with the wild-type T4 polymerase in the presence of dATP (50 μM) and d(AP)TP (285 μM). The direction of electrophoresis was from right to left. Only the lower part of the electrophoretogram is shown.

fairly follow the values measured in the original experiments (first two columns), except that they are increased by a factor of approximately 1.5, ascribable to the 1.5-fold higher input ratio of d(AP)TP/dATP used in these experiments.

A very marked skewing toward the shortest isostichs was noted in the 18-h hydrolysate of the fragment *Hin*fI-16b in which all nine interior A sites were exclusively occupied by (AP) (Figure 3b). However, in the fragments of actual interest, in which the A sites are only sparsely occupied by (AP), a hydrolysis of this duration does not produce this distorting effect. Thus, a comparison of the electrophoretic profiles obtained after 16- and 6-h hydrolyses of the fragments listed in Table II showed no difference.

The present results show that the extent of incorporation of (AP) for Ade at a given A site is not simply determined by the identity of the preceding base but is a much more complex function of the local base sequence. *E. coli* polymerase I, which incorporates (AP) at A sites with approximate indifference to sequence as judged by the nearest-neighbor criterion (Pless et al., 1981), synthesizes, in the form of its large proteolytic fragment (enzyme A), a ϕX174 *Hin*fI-16b fragment in which (AP) replaces Ade to the extent of 15% at position 381, 17% at position 377, and 9% at position 374 (Table I), even though all three of these sites are preceded by T. The variation of (AP) incorporation from one site to another is larger for the T4D enzyme, and larger yet in the case of the antimutator polymerase L141. For the latter enzyme, the percent incorporation of (AP) ranged from undetectable at some sites to 20% in position 2850 and 18% in position 2838. Since in these experiments dATP and d(AP)TP were competing at approximately equimolar input, and since with L141 polymerase the apparent K_m for insertion of d(AP)MP is about 6 times as large as the apparent K_m for insertion of dAMP (Clayton et al., 1979), the bias against (AP) incorporation at such sites as 2850 and 2838 may be wholly ascribable to this difference in K_m. Additional discrimination against (AP) incorporation by preferential exonucleolytic cleavage of 3′-terminal (AP) may be essentially absent at such sites.

For the L141 enzyme, Table III presents, in unified form, the values shown in Tables I and II for incorporation of (AP) at various A sites in different DNA fragments. The entries detail the tetranucleotide sequences preceding and succeeding the A site under consideration, and they are arranged first according to the identity of the immediately preceding base—in order to facilitate comparison with the nearest-neighbor results—and then, within this nearest-neighbor group, in order of decreasing (AP) incorporation at the A site. We observe that, whichever the preceding base may be, the incorporation of (AP) can still span the range from at least moderately high values to undetectable. Directing our attention to specific short sequences, we note that CC*A* often contains high levels of (AP)(14%, 14%, 14%, and 10% at positions 1171, 1131, 3406, and 1141, respectively), but we can also find intermediate (8%, position 2828) and low (4%, 4%, and 3% at positions 2820, 2833, and 2867, respectively) levels of (AP) incorporated at such sites. The sequence C*A*TTT is found associated with high incorporation of (AP) (11%, position 2857) and with low incorporation of (AP) (2% and −2% at positions 3440 and 3400, respectively). Finally, we find the identical heptanucleotide sequence, TTT*A*GTA, one time containing a high level of (AP) incorporation (14%, position 2853) and another time devoid of (AP) (−2%, position 3395). In this instance, differences are found only in positions removed by four nucleotide units upstream and downstream from the A site under consideration (ATTT*A*GTAG vs. TTTT*A*GTAA). Thus, we cannot on the basis of the data in Table III formulate any general principles which would allow a prediction of the extent of (AP) incorporation at an A site in a postulated sequence.

In our previous study by nearest-neighbor analysis (Pless et al., 1981), we found in experiments conducted at 10 and 20 °C, with both dATP and d(AP)TP present at 50 μM concentration, that stable incorporation of (AP) by L141 polymerase was relatively favored in positions following G and T and distinctly discriminated against if the preceding base was A. In the present work, we obtain (Table III) average values for (AP)/[A + (AP)] of 6.3% for C*A* sequences (averaged over 19 values), 2.9% for A*A* sequences (17 values), 3.5% for G*A* sequences (11 values), and 11.0% for T*A* sequences (10 values). Thus, here again, average incorporation of (AP) decreases in going from T*A* to C*A* to A*A* nearest-neighbor sequences. The low average value of (AP) incorporation following A sites can only to a minor extent be ascribed to the fact that some of the preceding A sites are occupied by (AP), as the level of (AP) incorporation was generally low; it must rather be interpreted as reflecting a significant bias against the A(AP) sequence. The unexpected result in this nearest-neighbor inspection of the data in Table III is the low average (AP) incorporation after G, which differs very markedly from our findings in the nearest-neighbor experiments. Two reasons may account for this difference. First, in view of the wide fluctuation of the values of (AP)/[A + (AP)], the data obtained in the present work are far too few to provide averages comparable in accuracy to the nearest-neighbor values, which issue from the sampling of a very large multiplicity of local sequences subsumed under the heading of one nearest-neighbor sequence. A second, more funda-

Table III: Relative Incorporation of (AP) by L141 Polymerase into Different Local Sequences[a]

nucleotide position	local sequence (5′→3′)	(AP)/[A + (AP)] (%)	nucleotide position	local sequence (5′→3′)	(AP)/[A + (AP)] (%)
1171	GGCC*A*CGAC	14	2842	GTAA*A*GTTA	7
1131	GTCC*A*TCTC	14	1140	GCCA*A*CGGC	6
3406	GTCC*A*GTTG	14	2843	GGTA*A*AGTT	5
2857	CGGC*A*TTTA	11	1163	CGCA*A*TGGA	4
1141	CGCC*A*ACGG	10	1155	AGAA*A*GACG	4
391	ACTC*A*TGAT	9	1156	GAGA*A*AGAC	3
2828	AACC*A*TGAA	8	2866	ACCA*A*TCCG	3
3425	CGGC*A*ATCT	7	2831	CCAA*A*CCAT	3
1164	ACGC*A*ATGG	6	2819	ACCA*A*CATA	3
2876	ATTC*A*GCGA	6	1182	GTCA*A*TAGC	2
1183	AGTC*A*ATAG	4	2832	ACCA*A*ACCA	2
2820	AACC*A*ACAT	4	2824	ATGA*A*ACCA	2
2833	GACC*A*AACC	4	2823	TGAA*A*CCAA	2
2867	AACC*A*ATCC	3	3424	GGCA*A*TCTC	2
3433	TTGC*A*TCTC	3	1176	AGCA*A*GGCC	1
1177	TAGC*A*AGGC	2	2870	CGAA*A*CCAA	1
3440	TCTC*A*TTTT	2	2871	GCGA*A*ACCA	0
2817	CAAC*A*TAAA	1			2.9[c]
3400	TTGC*A*TTTT	−2			
		6.3[b]			
1147	GGAG*A*GCGC	6	2850	TAGT*A*GCGG	20
1159	ATGG*A*GAAA	6	2838	AGTT*A*GACC	18
2836	TTAG*A*CCAA	6	2844	CGGT*A*AAGT	16
1149	ACGG*A*GAGC	5	374	TATT*A*GTGG	14
2825	CATG*A*AACC	4	2853	ATTT*A*GTAG	14
1153	AAAG*A*CGGA	3	381	TCTT*A*CCTA	11
1168	CACG*A*CGCA	3	2808	CGTT*A*TTGC	11
1157	GGAG*A*AAGA	2	1180	CAAT*A*GCAA	6
388	CATG*A*TTTC	2	377	ACCT*A*TTAG	2
3413	TCTG*A*TTGT	1	3395	TTTT*A*GTAA	−2
2872	AGCG*A*AACC	0			11.0[e]
		3.5[d]			

[a] Synthesized with d(AP)TP/dATP = 1.2. [b] Average overall C*A*. [c] Average overall A*A*. [d] Average overall G*A*. [e] Average overall T*A*.

mental reason may be that the DNA sampled in the present investigation was subjected to a selective rigor which was entirely absent in our earlier study: in the nearest-neighbor work, *all* radioactive (AP) nucleotides which had been incorporated into the DNA (i.e., which had become acid insoluble) were equally weighted in the analysis; in contrast, the present paper reports the levels of (AP) incorporation at specified A sites *realized in only those growing DNA chains which were elongated beyond a certain position* (viz., the next *Hin*fI restriction site downstream). This stricture could entail a considerable selectivity in the particular type of DNA fragments assayed, as the DNA polymerase extended only a minor fraction of the DNA primers to the requisite length.

Arranging the data for the L141 polymerase experiment in Table III according to the nucleotide immediately *succeeding* the A site of interest, one obtaines average values for (AP)/[A + (AP)] of 5.2% for *A*C sites (10 values, ranging from 1% to 14%), 4.4% for *A*A sites (17 values, 0–16%), 8.5% for *A*G sites (14 values, −2–20%), and 4.6% for *A*T sites (16 values, −2–14%). Thus, essentially the full range of incorporation probabilities can be observed for each of the four groups. The average of 8.5% (AP) incorporation calculated from the *A*G sequences clearly exceeds the average values obtained for the other sequences. This could indicate that a succeeding G residue confers on an incorporated (AP) a special measure of protection against exonucleolysis by "peel back" (Goodman et al., 1974).

The results show that local helix stability is not the only criterion governing the propensity for (AP) incorporation at a given A site. This fact is most evident from an examination of Table I: if there were only one important parameter (e.g., helix stability), one would expect all three enzymes to show the same biases for or against misincorporation of (AP) at the different A sites, although these preferences might be more pronounced for one polymerase than for another; however, Table I indicates that while such parallel biases exist for the L141 and T4D enzymes, the sequence preferences for (AP) incorporation by *E. coli* polymerase I (enzyme A) and by the T4 enzymes can be reversed; this is a clear indication that at least two potentially antagonistic effects are operative.

For various mutant T4 polymerases, Muzyczka et al. (1972) have demonstrated a clear correlation between replicative accuracy and activity of polymerase-associated 3′–5′-exonuclease. In vitro studies by Clayton et al. (1979) have shown that for a given T4 polymerase the ratio of polymerizing to exonucleolytic action can be manipulated by adjustment of the deoxynucleoside triphosphate concentration, with an attendant alteration in the accuracy of DNA replication as measured by incorporation of (AP) for Ade. Fersht & Knill-Jones (1981) have obtained evidence that the misincorporation of canonical nucleotides by *E. coli* polymerase III holoenzyme can depend on the concentration of the deoxynucleoside triphosphate which is incorporated *after* the mismatched nucleotide. Analogous results were obtained by Kunkel et al. (1981) for *E. coli* polymerase I and for wild-type T4 polymerase. Misincorporation of nucleotides appears to be a complex function of the propensity for the original misinsertion of the nucleotide, of the probability of immediate excision of the misinserted nucleotide, of the rapidity with which DNA polymerization ensues to the downstream side of the misinserted nucleotide, of the overall activity of polymerase-associated 3′–5′-exonuclease, and of the propensity of the enzyme for dissociation from the template-primer. All of these processes are potentially sequence dependent in their rate and extent. Indeed, the mutagenic action of (AP) at a given site can be strongly influenced by the identity of the nucleotides

in the immediate vicinity (Coulondre et al., 1979) and even by single base pair changes at a considerable distance from the site (Conkling et al., 1980). Insofar as *misincorporation* of (AP) in vitro is concerned, its strong sequence dependence is documented by our earlier nearest-neighbor work (Pless et al., 1981) and, at a more detailed level, in the present report.

Difficulties arise when one attempts to interpret the marked differences in (AP) incorporation from one A site to another in terms of simple physical principles. It stands to reason that the polymerase in its intimate association with the template-primer can profoundly affect the structure and dynamics of the DNA in ways which are not expressed in the behavior of protein-free duplex DNA. The sequence specificity of DNA–protein interactions is as yet poorly understood; thus, DNA gyrase (in the presence of oxolinic acid) cleaves double-stranded DNAs with high accuracy at defined sites, which, however, do not have any obvious sequence elements in common (Fisher et al., 1981). Also, it is difficult to assess what influence the accessory proteins involved in the T4 replication complex (Alberts et al., 1982) would have in vivo.

Even if a physical rationale for the observed sequence dependence of (AP) incorporation is lacking, the experimental results reported here are significant in that they document that misincorporation frequency can vary markedly from one defined DNA site to another. This does not imply that positions of high misincorporation of (AP) correlate directly with "hot spots" for (AP)-induced mutagenesis. Other critical elementary processes in the putative mechanism of (AP)-induced transitions could also exhibit a marked sequence dependence: Watanabe & Goodman (1981) have shown that the frequency of misincorporation of C for T opposite a template (AP) moiety by KB cell polymerase α is strongly modulated by the identity of the surrounding nucleotides. Nonetheless, the variability of (AP) incorporation from one DNA site to another is of such magnitude that it has to be considered a major determinant in the composite sequence dependence of (AP) mutagenesis.

Acknowledgments

We benefited from discussion with the following people: Richard P. Cunningham, Keith McKenney, Hamilton O. Smith, and Barbara Sollner-Webb. We thank Dorothy Regula for help in preparing the manuscript.

Registry No. DNA polymerase, 9012-90-2; AP, 452-06-2; d-(AP)TP, 26666-45-5; d-$(pT)_3$p(AP), 86885-27-0; d-$(pT)_3$[p(AP)]$_2$, 86900-64-3; [5′-^{32}P]d-$(pT)_3$p(AP), 86885-28-1; [5′-^{32}P]d-$(pT)_3$[p-(AP)]$_2$, 86900-65-4; [5′-^{32}P]d-$(pA)_4$, 86885-29-2.

References

Alberts, B. M., Barry, J., Bedinger, P., Formosa, C. V., Jongeneel, C. V., & Kreuzer, K. N. (1982) *Cold Spring Harbor Symp. Quant. Biol. 47*, 655–668.

Benbow, R. M., Hutchison, C. A., Fabricant, J. D., & Sinsheimer, R. L. (1971) *J. Virol. 7*, 549–558.

Bessman, M. J., Muzyczka, N., Goodman, M. F., & Schnaar, R. L. (1974) *J. Mol. Biol. 88*, 409–421.

Bollum, F. J., Chang, L. M. S., Tsiapalis, C. M., & Dorson, J. W. (1974) *Methods Enzymol. 29*, 70–81.

Burton, K., & Petersen, G. B. (1960) *Biochem. J. 75*, 17–27.

Clayton, L. K., Goodman, M. F., Branscomb, E. W., & Galas, D. J. (1979) *J. Biol. Chem. 254*, 1902–1912.

Conkling, M. A., Koch, R. E., & Drake, J. W. (1980) *J. Mol. Biol. 43*, 303–315.

Coulondre, C., Miller, J. H., Farabaugh, P. J., & Gilbert, W. (1979) *Nature (London) 274*, 775–780.

Cunningham, R. P., DasGupta, C., Shibata, T., & Radding, C. M. (1980) *Cell (Cambridge, Mass.) 20*, 223–235.

Danna, K. J. (1980) *Methods Enzymol. 65*, 449–467.

Fersht, A. R., & Knill-Jones, J. W. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4251–4255.

Fisher, L. M., Mizuuchi, K., O'Dea, M. H., Ohmori, H., & Gellert, M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4165–4169.

Freese, E. (1959) *J. Mol. Biol. 1*, 87–105.

Glynn, I. M., & Chappell, J. B. (1964) *Biochem. J. 90*, 147–149.

Goodman, M. F., Gore, W. C., Muzyczka, N., & Bessman, M. J. (1974) *J. Mol. Biol. 88*, 423–435.

Klenow, H., Overgaard-Hansen, K., & Patkar, S. A. (1971) *Eur. J. Biochem. 22*, 371–381.

Koch, R. E. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 773–776.

Kunkel, T. A., Schaaper, R. M., Beckman, R. A., & Loeb, L. A. (1981) *J. Biol. Chem. 256*, 9883–9889.

Lo, K. Y., & Bessman, M. J. (1976) *J. Biol. Chem. 251*, 2475–2479.

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Muzyczka, N., Poland, R. L., & Bessman, M. J. (1972) *J. Biol. Chem. 247*, 7116–7122.

Pless, R. C., Levitt, L. M., & Bessman, M. J. (1981) *Biochemistry 20*, 6235–6244.

Ronen, A., Rahat, A., & Halevy, C. (1976) *Genetics 84*, 423–436.

Ronen, A., Halevy, C., & Kass, N. (1978) *Genetics 90*, 647–657.

Salts, Y., & Ronen, A. (1971) *Mutat. Res. 13*, 109–113.

Sanger, F., Coulson, A. R., Friedman, 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.

Setlow, P., Brutlag, D., & Kornberg, A. (1972) *J. Biol. Chem. 247*, 224–231.

Swanstrom, R., & Shank, P. R. (1978) *Anal. Biochem. 86*, 184–192.

Watanabe, S. M., & Goodman, M. F. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2864–2868.

Weiss, B., Live, T. R., & Richardson, C. C. (1968) *J. Biol. Chem. 243*, 4530–4542.

Weymouth, L. A., & Loeb, L. A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 1924–1928.