Misincorporation during DNA synthesis, analyzed by gel electrophoresis

Greg G.Hillebrand⁺, Alan H.McCluskey, Katherine A.Abbott, Galina G.Revich and Kenneth L.Beattie*

Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, TX 77030, USA

Received 3 January 1984; Revised and Accepted 12 March 1984

ABSTRACT

A method has been developed for simultaneous comparison of the propensity of a DNA polymerase to misincorporate at different points on a natural template-primer. In this method elongation of a [$5'$ - 32 P]primer, annealed to a bacteriophage template strand, is carried out in the presence of only three dNTPs (highly purified by HPLC). Under these conditions the rate of primer elongation (monitored by gel electrophoresis/autoradiography) is limited by the rate of misincorporation at template positions complementary to the missing dNTP. Variations in the rate of elongation (revealed by autoradiographic banding patterns) reflect variations in the propensity for misincorporation at different positions along the template. The effect on primer elongation produced by addition of a chemically modified dNTP to 'mimic' reactions reveals the mispairing potential of the modified nucleotide during DNA synthesis. By use of this electrophoretic assay of misincorporation we have demonstrated that the fidelity of *E. coli* DNA polymerase I varies greatly at different positions along a natural template, and that BrdUTP and IodUTP can be incorporated in place of dCTP during chain elongation catalyzed by this enzyme.

INTRODUCTION

Several methods have been used to investigate the accuracy of DNA synthesis with purified DNA polymerases. Hall and Lehman (1) used a procedure developed by Trautner *et al.* (2) to quantitate the fidelity of DNA synthesis. In this approach, which is still widely used, a synthetic template-primer of defined repeating base sequence [e.g., poly(dA-dT)] is used to measure incorporation of a noncomplementary nucleotide (e.g., [α - 32 P]dGTP) along with a complementary nucleotide (e.g., [3 H]dTTP). The ratio of [32 P] to [3 H] incorporated into acid insoluble material in such an experiment provides a quantitative estimate of the misincorporation rate under a given set of conditions. In a second approach, developed in the laboratory of Loeb (3,4), a natural DNA template-primer is used, consisting of a mutant viral strand of bacteriophage ϕ X174, to which an appropriate restriction fragment primer is annealed. Following chain elongation past the mutant site on the template, the DNA is used in a transfection assay. The frequency of revertant plaques produced from the *in vitro*-synthesized DNA is used to estimate the error frequency during the polymerization. In a third technique (5-7), the fidelity of DNA synthesis is assessed by measuring the competition between 2-aminopurine deoxynucleoside triphosphate and deoxyadenosine triphosphate for incorporation at thymine sites on natural or synthetic DNA templates.

The above techniques have been very valuable in defining the influence of various factors on the fidelity of DNA synthesis [recently reviewed by Loeb & Kunkel (8)]. These include cellular source of polymerase, enzymatic properties such as subunit makeup and

the presence of associated 3'-exonuclease activity, and external factors such as the presence of DNA binding proteins and mutagenic divalent metal ions. In this paper we introduce a new experimental approach for the analysis of misincorporation during DNA synthesis, one whose major contribution is the ability to examine the relative accuracy of polymerization at many different positions on a natural DNA template. As demonstrated in this and the accompanying paper (Hillebrand & Beattie), the nucleotide sequence of the template exerts a striking influence on the frequency of mispairing during DNA synthesis.

MATERIALS AND METHODS

DNA, Enzymes and Chemicals

DNA of bacteriophage G4 (viral strand) was prepared as described previously (9), except for the following modifications. Cells were infected in medium containing 2mM potassium phosphate (pH 7) and no [32 P]orthophosphoric acid. To achieve more reproducible lysis than obtained by the previous method, cells were harvested at 1 hr after infection, then incubated in an equivalent volume of warm medium lacking sucrose. DNA from bacteriophage M13 was prepared as described by Messing (10). The covalently closed replicative form I DNA (RFI) of phage G4 was labeled with [3 H] and isolated as described by Cunningham et al. (11). Synthetic oligonucleotide primers for use with M13 templates were purchased from P-L Biochemicals. DNA polymerase I 'large fragment' of *E. coli* was purchased from New England Nuclear, Bethesda Research Laboratories, Boehringer Mannheim and P-L Biochemicals. Although all data presented in this and the accompanying paper were obtained with commercial sources of *E. coli* pol I, we have recently begun using enzyme purified in our own laboratory for the electrophoretic assay of misincorporation, because of a recurring problem with commercial enzymes: most lots of the 'large fragment' pol I (from all suppliers listed above) contain a contaminating activity (5'-exonuclease or phosphatase) that removes the [$5'$ - 32 P] label from primers during polymerase reaction. In some, but not all cases, this activity can be inhibited by including 200mM KCl in the reaction. Restriction endonucleases were obtained from Bethesda Research Laboratories, New England Biolabs and Boehringer Mannheim. Calf intestine alkaline phosphatase and polynucleotide kinase were obtained from Boehringer Mannheim. 5-Bromo-2'-deoxyuridine 5'-triphosphate (BrdUTP) and 5-iodo-2'-deoxyuridine 5'-triphosphate (IodUTP) were purchased from P-L Biochemicals. HPLC-purified 2'-deoxynucleoside 5'-triphosphates (dNTPs) were purchased from ICN Pharmaceuticals, Inc. and from P-L Biochemicals. All dNTPs were further purified as described below. The 2',3'-dideoxynucleoside 5'-triphosphates (ddNTPs) were purchased from P-L Biochemicals. [3 H]Thymidine (40-60 Ci/mmol) was from Schwarz-Mann. [32 P]ATP (>5000 Ci/mmol) and [α - 32 P]dNTPs (3000 Ci/mmol) were obtained from the Amersham Corp. Sepharose 2B and DEAE-Sepharose were supplied by Pharmacia Fine Chemicals, and DE81 filter disks were purchased from Whatman. All other chemicals were of reagent grade quality.

Preparation of template-primer DNA

Approximately 250 μ g of [3 H]G4 RFI of known specific activity was cleaved with the desired restriction endonuclease according to the conditions recommended by the supplier. Reactions were stopped by addition of excess EDTA (over divalent cation) and the resulting double-stranded fragments were separated by electrophoresis at 200V through a 5% polyacrylamide slab gel (16 cm x 24 cm x 1.5 mm) with a 15% polyacrylamide segment at the bottom (16 cm x 12 cm x 1.5 mm). Electrophoresis buffer consisted of 40 mM Tris base, 20mM sodium acetate and 1mM Na $_2$ EDTA, adjusted to pH 7.2 with HCl. After appropriate time of electrophoresis, the gel was stained with 2.5 μ g/ml ethidium bromide and fragments were visualized by illumination with long-wavelength ultraviolet light. Alternatively, if fragments were previously 5'-labeled with [32 P], they were located by autoradiography. The bands were excised from the gel and the DNA was extracted by homogenization with a Tekmar Tissue mixer in 1-2 ml of 10mM Tris-HCl (pH 7.5), 1mM Na $_2$ EDTA (TE buffer). After the mixture stood overnight at 4°C or 3 hr at room temperature, the polyacrylamide was removed by filtration through a Bio-Rad Econo-Column (1 cm x 10 cm with fritted glass disk). The column was washed with 1-2 ml TE buffer, and remaining liquid was forced through with air pressure. DNA was precipitated from combined eluent (collected in 5/8-inch x 4-inch polyallomer tube) by addition of 1/10 volume 3M sodium acetate and 2 volumes cold ethanol. After overnight at -20°C or 30 min in ethanol-dry ice, tubes were centrifuged at 25,000 rpm in a Beckman SW27.1 rotor for 30 min at 4°C. Precipitated DNA was rinsed with cold 70% ethanol, dried under vacuum, then

Table 1. Primers Used in This Study

Primer Designation	Enzymes Used For Generation	Length (Residues)	3'-OH Termin ^a of Primer
F66	HinfI	66	2720
A22	AluI	22	4070
RA50	RsaI, AluI	50	1062
RA69	RsaI, AluI	69	1112
S17(M13mp9)	(Synthetic)	17	6254(M13mp9)

^a Template residue at which first incorporation occurs, according to nucleotide sequence of bacteriophage G4 (12) or M13mp9 (10,13).

dissolved in a small volume (usually 100 μ l) of TE buffer. DNA concentrations were determined by measurement of radioactivity, using previously determined specific activities. Fragments, obtained in 50-80% yield (starting from undigested DNA), were stored at -20°C . Table 1 lists the primers used in the electrophoretic assays of misincorporation presented in this and the accompanying paper. Many additional primers (not listed in the table) were used during this investigation to generate data that are commented upon in the text or represented in sequence listings (Figure 3 of this paper and Figure 4 of the accompanying paper). Nucleotide sequences referred to in Table I and displayed in the figures were determined by Godson *et al.* (12) for phage G4 and by van Wezenbeek *et al.* (13) and Messing (10) for phage M13.

For end-labeling of individual restriction fragments, a quantity of DNA corresponding to 2-20 pmol 5'-end was extracted three times with an equal volume of isoamyl alcohol saturated with TE buffer to remove any remaining ethidium bromide. The DNA was then precipitated with ethanol as described above, dried under vacuum and dissolved in 20 μ l TE buffer. When an entire restriction digest was to be end-labeled, after termination of digestion with EDTA, the DNA was ethanol precipitated, then dissolved in 20 μ l TE buffer for the subsequent phosphatase reaction. The 5'-terminal phosphoryl groups were removed by treatment with calf intestine alkaline phosphatase in a 50- μ l reaction mixture containing 200mM Tris-HCl (pH 8.6), 1mM MgCl_2 , 0.15mM zinc acetate and approximately 7 units of phosphatase. After 30 min at 37°C , the reaction was terminated by adding EDTA (2mM) and heating at 65°C for 15 min, then the DNA was extracted three times with phenol saturated with TE buffer, precipitated with ethanol, dried under vacuum, then dissolved in 20 μ l TE buffer. Isolated fragments were labeled with [^{32}P] at their 5'-termini in a reaction mixture (25-50 μ l) containing 70mM Tris-HCl (pH 8), 10mM MgCl_2 , 5mM DTT, 20-200 μCi [^{32}P]ATP and 5-25 units of polynucleotide kinase. To determine the specific activity of the [^{32}P] fragments after the kinase reaction, 1- μ l aliquots of the reaction mixture, both before addition of kinase and after 1-2 hr reaction at 37°C , were applied to Whatman DE81 disks. These were processed as described by Rowen & Kornberg (14), dried, and placed in scintillation vials with 5 ml toluene-PP0-POPOP scintillation fluor. Quantitation of [^{32}P] was done with a Beckman LS-233 liquid scintillation counter. The [^{32}P] radioactivity (after subtraction of counts on first disk), along with the known concentration of DNA in the kinase reaction, was used to calculate the specific activity of [5'- ^{32}P] fragment. After isolated fragments were 5'-end labeled, they were precipitated with ethanol to remove most of the unincorporated [^{32}P]ATP, and dissolved in 20 μ l TE buffer.

To prepare a G4 template-primer a 5-fold molar excess of template, 400mM NaCl and 50mM Tris-HCl (pH 8) were added to the [5'- ^{32}P] fragment, then the mixture (50-100 μ l) was heated at 100°C for 3 min, then at $55-60^{\circ}\text{C}$ for 1 hr. To remove unincorporated [^{32}P]ATP and annealed fragments, the mixture was passed over a Sepharose 2B column (0.7 cm x 20 cm). Elution was carried out with TE buffer and 4-drop fractions were collected. The void volume peak containing template-primers (located by Cerenkov counting) was pooled and the DNA was precipitated with ethanol and dissolved in 100 μ l TE buffer. A small aliquot was counted for [^{32}P] to determine the primer concentration. When primers consisted of chemically synthesized oligonucleotides (complementary to M13 viral DNA) the above protocol was used, except that the oligonucleotide was [5'- ^{32}P]-labeled without necessity for prior treatment with phosphatase, and annealed with only a slight molar excess of template. Template-primers were stored at -20°C until used.

Purification of dNTPs by HPLC

'HPLC-pure' dNTPs from ICN Pharmaceuticals or from P-L Biochemicals were further purified by high performance liquid chromatography (at room temperature) on a Dupont Zorbax SAX column or a Whatman Partisil SAX column (25 cm x 4.6 mm). The latter column was packed by Custom LC, Inc. (Houston, TX). With the Dupont column baseline separation of the 4 dNTPs was achieved by isocratic elution at 1.5 ml min⁻¹ with 500mM KH₂PO₄, pH 3.1. With the Whatman column complete resolution of the 4 dNTPs was attained by isocratic elution at 2.0 ml min⁻¹ with 750mM KH₂PO₄, pH 3.45. The order of elution of dNTPs was C,A,T,G with the Dupont column, and C,T,A,G with the Whatman column. For purification of any given dNTP, conditions were used that permitted the greatest separation of that dNTP from the others, regardless of whether any contaminating dNTPs were separated from each other. In the purification of nucleotide analogs, conditions were chosen to give optimal separation from any possible contaminating dNTP that could cause artifactual chain elongation in the electrophoretic assay of misincorporation. Mobile phase was delivered with a Perkin-Elmer Series 4 Liquid Chromatograph and ultraviolet absorption data were collected with a Hewlett Packard HP1040A spectrophotometric detection system. To efficiently monitor all nucleotide species, a peak wavelength of 260 nm with 80 nm bandwidth was chosen as absorption signal to be stored on disk during the chromatography. The diode array detector was set to take absorption spectra from 205-400 nm at 2-nm intervals; spectra were taken automatically at the upslope, apex and downslope of each peak, and stored on disk. Post-run analyses and plotting were performed with resident system software, utilizing the HP-85 microcomputer and HP7470A plotter.

For preparative purifications, up to 10 µmol dNTP solution (in water or mobile phase) was loaded onto the column and 0.500-ml fractions were collected. Peak fractions could be directly reinjected onto the column to provide further purification. Pooled fractions were desalted by DEAE-Sephacel chromatography with the volatile buffer, triethylammonium bicarbonate (15). Nucleotide concentrations were determined spectrophotometrically with a Zeiss model PM6 spectrophotometer using the extinction coefficients supplied by ICN or P-L Biochemicals.

Electrophoretic Assay of Misincorporation

The accuracy of DNA synthesis was assayed by allowing the DNA polymerase to elongate the [5'-³²P]primer in the presence of only three dNTPs ('minus reaction'). Under these conditions the primer is rapidly elongated to the point on the template where the missing nucleotide would normally be incorporated. Further elongation requires the incorporation of a noncomplementary nucleotide at this point, after which elongation can proceed to the next point where the missing nucleotide is needed, etc. The extent of primer elongation, which reflects the extent of misincorporation, was analyzed by gel electrophoresis as follows. Polymerization was stopped by addition of excess EDTA, and the DNA was precipitated with ethanol, then dissolved in 2-5 µl of deionized formamide containing 0.1% each of bromophenol blue and xylene cyanol FF. DNA was denatured at 100°C for 3 min, then applied to an 8-10% polyacrylamide gel (16 cm x 37 cm x 0.4-1.5 mm) containing 7M urea and TBE buffer (90mM Tris-borate, 2.5mM Na₂EDTA, pH 8.3). Electrophoresis was carried out at 30 mA for a time dependent on the initial primer length, then one of the glass plates was removed, and the gel was covered with Saran Wrap and autoradiographed using Kodak XR-1, XAR-5 or Fuji RX X-ray film. In some experiments a Dupont Cronex Hi-Plus intensifying screen was used during autoradiography. Details of the reaction conditions are given in the figure legends.

RESULTS

Effect of Successive HPLC Purification of dNTPs on the Extent of Chain Elongation in 'Minus' Reactions

To determine whether commercially available dNTPs are sufficiently pure to give valid results in the electrophoretic assay of misincorporation, we compared the results obtained in assays performed with commercially available 'HPLC-pure' dNTPs with those obtained after the dNTPs (obtained from P-L Biochemicals) were further purified, by a second and third cycle of HPLC (see Materials and Methods). The conditions chosen to further purify the commercial dNTP preparations permitted removal of deamination products (dUTP

and dTTP) that could be formed spontaneously during storage of dCTP and dATP, respectively. These deamination products, if present in the 'minus' reactions, would create problems, since dUTP could replace dTTP in chain elongation, and dITP could replace dGTP.

As will be reported elsewhere (Revich & Beattie, in preparation), during repurification of all of the commercial HPLC-purified dNTPs, several contaminating peaks were seen, but they nearly all had UV absorption spectra that resembled those of the main peaks. This suggests that the major source of 'contamination' is spontaneous hydrolysis and disproportionation reactions occurring during storage, yielding free bases and nucleosides, mono- and diphosphates, and tetraphosphate species (the latter eluting after dNTP). In no case was there any evidence of the presence of deamination products in the dNTP preparations. In only one case was there indication of minor cross-contamination of dNTP species. During the repurification of dATP we detected a very small peak whose retention time and absorption spectrum were as expected for dTTP. From spectrophotometric data we estimate that the level of contamination of the commercially purified dATP by dTTP was on the order of 0.1%.

Figure 1 shows the results of the electrophoretic assay of fidelity of *E. coli* pol I, conducted with successively purified dNTPs. The electrophoretic mobility of the unelongated 17-mer primer is shown in lane 1. Lanes 2, 6, 10 and 14 represent 'dideoxy' reactions (ddA, ddG, ddC and ddT, respectively). Each 'dideoxy' reaction generated a collection of elongated primers, terminated by incorporation of the 2',3'-dideoxy analog at positions on the template at which the corresponding dNMP would normally be incorporated. 'Dideoxy' bands in the autoradiograph mark the positions on the template at which misincorporation may occur in the corresponding 'minus' reaction. Lines connect each 'ddA' band in Figure 1 with its complementary residue in the M13mp9 template.

The lack of significant differences in any given 'minus' reaction conducted with dNTPs purified by HPLC once (commercial preparation), twice or three times confirms that our results are not due to cross-contamination of the dNTPs. Furthermore, the HPLC-purified dNTPs obtained from P-L Biochemicals may be adequately pure for use in the electrophoretic assay of misincorporation. The latter point may not apply for all commercially purified preparations of dNTPs, and certainly does not apply for all template-primers. In the region of the M13mp9 template used for analysis of misincorporation in Figure 1, there was a relatively large amount of misincorporation in place of dTTP (lanes 15-17), compared with that observed under similar conditions using many primers. This may account for the barely detectable decrease in chain elongation in the '-T' reaction that accompanied removal of the presumed contaminating dTTP from the preparation of dATP (compare lanes 15-17). Experiments in which we intentionally added various concentrations of contaminating dNTPs to 'minus' reactions (data not shown) indicated that with certain template-primers, increased chain elongation occurred upon addition of contaminating dNTP at a level of one part in 10,000 (approximately ten times lower than the presumed contamination in the '-T' reaction of Figure 3). Therefore, we stress the importance of further purification of any commercial preparation of dNTP prior to its use in the electrophoretic assay of misincorporation.

The results shown in Figure 1 provide strong evidence for the validity of our assay as a method for detection of misincorporation. Furthermore, the fact that site-specific mutagenesis has been achieved by a gap-filling polymerase reaction carried out in the absence of one dNTP (16-18) confirms that misincorporation occurs in reactions

Figure 1. Effect of Successive HPLC Purification of dNTPs on Elongation of Primer S17 in the Presence of Three dNTPs. Reaction mixtures (25 μ l total volume) contained 32mM Tris-HCl (pH 7.5), 5mM MgCl₂, 50 μ M dNTPs, 4nM (3'-OH termini) S17 template-primer and 40 units/ml *E. coli* DNA polymerase I ('large fragment'). After incubation at 30°C for 30 min, reactions were stopped by addition of excess EDTA. DNA was analyzed by electrophoresis in a 10% polyacrylamide, 7M urea gel, followed by autoradiography (See Materials and Methods). The template sequence is displayed on the left, with lines connecting T residues on the template with corresponding 'ddA' bands in the autoradiograph. Lane 1, unelongated 17-mer. Lanes 2, 6, 10 and 14, ddA, ddG, ddC and ddT, respectively. Lanes 3, 4 and 5, '-A' reactions containing dNTPs purified by 0, 1 and 2 cycles of HPLC, respectively, starting with commercially HPLC-purified material. Lanes 7, 8 and 9 represent the equivalent '-G' reactions, Lanes 11, 12 and 13 represent the '-C' reactions, and Lanes 15, 16 and 17 represent the '-T' reactions.

containing only three dNTPs. In addition, some features of the data obtained in electrophoretic analyses of chain elongation (pointed out below) are difficult to explain on the basis of incorporation of contaminating nucleotides. Finally, our analysis of the identity of nucleotides misincorporated in 'mimms' reactions (to be reported in a subsequent publication) directly proves that noncomplementary nucleotides are incorporated in place of the missing dNTP.

Sequence-dependent Phenomena in Misincorporation

The data of Figure 1 illustrate several characteristics of misincorporation that exhibit positional effects along the template. First, in comparing the bands produced in 'mimms' reactions with the corresponding 'dideoxy' bands, one can see that the points on the template at which chain elongation is rate-limiting (i.e., at which primers accumulate to produce a band in the autoradiograph) are not exclusively the positions at which the missing dNTP is needed. Additional bands are produced in 'mimms' reactions, corresponding to positions at which, following misincorporation, the rate of subsequent elongation is slow enough to produce a band with identical electrophoretic mobility as the 'dideoxy' band. As illustrated in the '-A' reaction, this phenomenon is seen at only certain positions on the template. In this reaction elongation occurred first up to a position on the template requiring two consecutive misincorporations before subsequent elongation could take place (the TT doublet near the 3'-OH of the initial primer). The '-A' reaction produced two bands in this region of the autoradiograph, as expected if elongation paused until misincorporation occurred at each of the two T residues on the template, then could resume immediately after insertion of two incorrect nucleotides. Thus, for each of these two 'ddA' bands there was one corresponding '-A' band, migrating at the expected position for a chain one residue shorter. At the next T residue on the template, chain elongation in the absence of dATP was retarded not only prior to misincorporation, but also after insertion of the incorrect nucleotide, resulting in a doublet in the '-A' lane, whose upper member had electrophoretic mobility identical to that of the 'ddA' band. Thus, this T residue in the template represents a 'postmisincorporation pause' site. At the next T on the template a single faint band was produced (no postmisincorporation pause), whereas at the next T on the template three bands appeared in the '-A' reaction, indicating that insertion of the two dNMP residues following misincorporation at this position was rate-limiting. Comparison of bands produced in '-T' reactions (lanes 15-17) with the template sequence also reveals that postmisincorporation pause occurs at some positions on the template and not at others. The existence of these extra 'mimms' bands is another feature of chain elongation in the presence of 3 of the 4 dNTPs that is inconsistent with chain elongation being the result of slow incorporation of contaminating dNTPs. On the other hand, the phenomenon is consistent with the known slow rate of elongation of primers containing mismatched 3'-termini (19,20). A feature of the sequence that appears to be associated with this phenomenon will be discussed later.

Another phenomenon that is revealed by the data of Figure 1 is variation in the relative rates of misincorporation at different positions along the template. Within a given 'mimms' lane this phenomenon is reflected by variation in band intensities; faint or missing bands in a 'mimms' lane (at the positions expected from the corresponding 'dideoxy' lane), flanked by dark autoradiographic bands, mark the positions of 'hotspots' for misincorporation. For example, in the '-A' lanes of Figure 1, the band corresponding to the fourth misincorporation site is very faint, compared to the bands corresponding to the

Figure 2. Electrophoretic Analysis of Elongation of Primer RA50 in the Presence of Three dNTPs. The template sequence near the 3'-end of the primer is displayed along the left edge, with lines connecting T residues in the template with corresponding 'ddA' bands in the autoradiograph.

Panel A. Effect of addition of heat-inactivated DNA polymerase on the extent of chain elongation catalyzed by *E. coli* DNA polymerase I in the presence of 3 dNTPs. Reaction mixtures (25 μ l total volume) contained 32mM Tris-HCl (pH 7.5), 5mM MgCl₂,

50 μ M dNTPs, 2 nM (3'-OH termini) RA50 template-primer and 26 units/ml E. coli DNA polymerase I ('large fragment'). Some reactions (indicated below) also contained an equivalent quantity of enzyme that had been inactivated by heating at 100°C for 10 min. Polymerization was carried out at 30°C for 30 min. Lane 1, unelongated RA50 primer. Lanes 2, 5, 8 and 11, ddA, ddG, ddC and ddT, respectively. Lanes 3 and 4, '-A' reactions without and with heat-inactivated enzyme, respectively. Lanes 6 and 7, '-G' reactions without and with heat-inactivated enzyme, respectively. Lanes 9 and 10, '-C' reactions without and with heat-inactivated enzyme, respectively. Lanes 12 and 13, '-T' reactions without and with heat-inactivated enzyme, respectively. Lane 14, primer elongated for 1 min in presence of all 4 dNTPs. Lane 15, same as lane 14, except dNTPs were heated at 100°C for 10 min before addition to polymerase reaction mixture. Lane 16, reaction mixture containing all 4 dNTPs, incubated for 1 min with only heat-inactivated polymerase.

Panel B, Timecourse of elongation of RA50 primer by E. coli DNA polymerase I in the presence of 3 dNTPs. Reaction mixtures contained 32 mM Tris-HCl (pH 7.5), 5 mM MgCl₂, 10 μ M dNTPs, 0.4 nM (3'-OH termini) RA50 template-primer and 40 units/ml DNA polymerase I ('large fragment') of E. coli. Lanes 1-4, '-T' reactions carried out for 5, 30, 60 and 180 min, respectively. Lane 5, ddT. Lanes 6-9, '-G' reactions carried out for 5, 30, 60 and 180 min, respectively. Lane 10, ddG.

third and fifth misincorporation sites, suggesting that misincorporation at the fourth T residue on the template occurred much more readily than that at the preceding and subsequent T residues on the template. More examples of this phenomenon will be shown below.

Also seen in Figure 1 is a mobility shift phenomenon associated with misincorporation at some sites. In the '-C' lanes (11-13) there are bands whose migration differs slightly from that expected from the adjacent 'ddC' bands (lane 10). The band corresponding to the primers elongated up to the first G on the template (which had not yet undergone misincorporation) migrated to the expected position (aligning with the 'ddT' band to the right). However, the subsequent '-C' bands migrated slightly more slowly than expected, in comparison with the 'ddC' bands in lane 10 (and also in comparison with the '-G' bands in lanes 7-9). This mobility shift appeared to translate further up the gel, as well. It is known that the electrophoretic mobility of oligonucleotides is affected by nucleotide sequence, with relative mobility in the order, C>A>T>G (21). Accordingly, in electrophoretic analyses of chain elongation in the presence of all 4 dNTPs, using a short primer, annealed to M13 templates differing in sequence at one position, we have noted that a slight decrease in electrophoretic mobility occurs upon incorporation of a G residue, as opposed to a C residue (data not shown). Although mobility shift appears to occur most commonly when misincorporation occurs in place of dCTP, it is not restricted to '-C' reactions. Furthermore, the phenomenon does not occur equally at all G residues in the template, which suggests an influence of template sequence on the type of mispairing that can occur during DNA synthesis. The mobility shift phenomenon is another indication that chain elongation in 'minus' reactions is not the result of slow incorporation of contaminating dNTPs.

The data of Figure 2, obtained with G4 primer RA50, further document the variation in the propensity for misincorporation at different positions along natural DNA templates. In panel A arrows point to two positions on the autoradiograph where bands in the 'minus' lanes are very faint or missing, compared with surrounding bands (corresponding to positions on the template at which misincorporation in place of dATP or dCTP occurs at relatively high frequency). The original purpose of the experiment of Figure 2A was to determine whether any of the elongation observed in 'minus' reactions could result from the

presence of trace dNTP contamination in a DNA polymerase preparation. For each of the four 'mimms' reactions, a parallel reaction was carried out, in the presence of excess heat-inactivated enzyme. Lane 16 shows that the enzyme was completely heat-inactivated (compare with lane 1, no enzyme). Lanes 14 and 15 show that heating did not hydrolyze the dNTPs. The identical patterns of autoradiographic bands produced from 'mimms' reactions performed in the presence and absence of heat-inactivated enzyme proves that dNTPs do not contaminate the DNA polymerase preparation. The procedure used to purify the DNAs (see Materials and Methods) would preclude any dNTP contamination arising from this component of the polymerase reactions.

The autoradiographic data of Figure 2B represent a timecourse of chain elongation in '-T' and '-G' reactions with the RASO template-primer (refer to template sequence to left of panel A). By use of longer reaction times and higher polymerase/primer ratio, we were able to examine the propensity for misincorporation throughout a longer segment of the template than in the experiment of panel A. The arrows marked 'a' point to 'dideoxy' bands whose corresponding 'mimms' bands are either missing or are very faint, compared with surrounding bands, indicating the existence of 'hotspots' for misincorporation at these sites. In the case of the '-T' reaction, the A residue in the template corresponding to arrow 'a' is 78 residues away from the 3'-terminus of the primer. Thus, at least with certain primers, the electrophoretic assay can be used to examine the fidelity of synthesis over a relatively long stretch of template.

We have observed that two or more consecutive misincorporations generally (but not always) occur much less frequently than single misincorporations. However, at template sequences requiring multiple misincorporations for further elongation in 'mimms' reactions, the last misinsertion often occurs much more readily than the first one. Thus, the 5'-member of a string of identical template residues often represents a 'hotspot' for misincorporation. An example of this occurrence is represented by the '-T' data of Figure 2B. Arrow 'b' points to a 'ddT' band in lane 5, which represents the last member of an AAA sequence on the template. Note that the '-T' band (lanes 1-4) at the position just below it, which represents primers that had undergone misincorporation at the first two members of the AAA sequence, steadily decreased in intensity with time, compared with the bands representing the first two members of the triplet. In the '-G' reaction (represented by lanes 6-9 of Figure 4B) this phenomenon occurred, though only evident at the latest time of reaction, at the third member of a CCC sequence on the template (see 'ddG' band pointed to by the upper arrow 'b', lane 10). The lower arrow 'b' in the right margin of panel B points to a 'ddG' band that appears to represent another 'warmspot' for misincorporation (which produces a 'mimms' band that steadily decreases in intensity, compared with bands above and below it).

A final point that can be made from the timecourse of misincorporation represented in Figure 2B concerns 'postmisincorporation pause' sites. The data further demonstrate that at certain template positions elongation following misincorporation occurs with considerable delay, giving rise to 'mimms' bands of electrophoretic mobility equal to that of the corresponding 'dideoxy' bands. A particularly severe case of this phenomenon occurs at the A residue in the template closest to the 3'-OH of the RASO primer. Many of the primers that underwent misincorporation at this position were quickly elongated further (as indicated by the production of bands further up the gel), but a significant fraction of primers that had undergone misincorporation at this position were very poorly elongated

Figure 3. Nucleotide Sequence Surrounding 'Postmisincorporation Pause Sites' Observed with DNA Polymerase I of *E. coli*. The diagram at the top represents a template, paired with a primer that has been elongated in a 'minus' reaction up to a template position (shaded) complementary to the missing dNTP. Nucleotide sequences are listed in the 5'-3' direction (left to right). In each sequence the central (boldface) residue represents a position on the template at which, following incorporation of a noncomplementary nucleotide in a 'minus' reaction, subsequent elongation and 3'-exonuclease action occur relatively slowly. These sites were identified by use of the electrophoretic assay of misincorporation with various G4 and M13 template-primers. Lower Panel, Occurrence of Specific Dinucleotide Pairs on the 3'-Side of 'Post-misincorporation Pause Sites', as a Function of the Relative Energy of Each Dinucleotide Pair. The data represent dinucleotides immediately to the right of the central (boldface) template residues listed at the top. Relative free energy values were as reported by Ornstein and Fresco (22).

further, even at the longest time of incubation.

A listing of the template sequence surrounding 46 'postmisincorporation pause sites' that we have observed with *E. coli* DNA polymerase I appears in the upper part of Figure 3. Although no obvious trend in nucleotide composition or sequence exists on the 5'-side

Figure 4. Electrophoretic Analysis of the Mispairing Potential of BrdUTP and IodUTP During Elongation of Primer RA69 by DNA Polymerase I of *E. coli*. Polymerization was carried out with 10 μ M unmodified dNTPs, as described in Figure 2B. Lane 1, ddA. Lanes 2-6, '-A' with no analog (lane 2), 10 μ M (lane 3) and 50 μ M (lane 4) BrdUTP, or 10 μ M (lane 5) and 50 μ M (lane 6) IodUTP. Lane 7, ddG. Lanes 8-12, '-G' with no analog (lane 8), 10 μ M (lane 9) and 50 μ M (lane 10) BrdUTP, or 10 μ M (lane 11) and 50 μ M (lane 12) IodUTP. Lane 13, ddC. Lanes 14-18, '-C' with no analog (lane 14), 10 μ M (lane 15) and 50 μ M (lane 16) BrdUTP, or 10 μ M (lane 17) and 50 μ M (lane 18) IodUTP.

of template residues at which elongation occurs slowly subsequent to misincorporation, on the immediate 3'-side of these template residues, there is a high tendency towards strong base stacking interactions. Using the data of Ornstein & Fresco (22) we have arranged the occurrence of these 'postmisincorporation pause sites' according to the relative energy of their 3' dinucleotide neighbors. As seen in the lower part of Figure 3, there is a strong correlation between the occurrence of this phenomenon and the relative stability of the double helix imparted by the dinucleotide pair on the 3'-side of the misincorporation sites.

Figure 4 illustrates the use of the electrophoretic assay of misincorporation to analyze the mispairing potential of chemically modified dNTPs during DNA synthesis. As seen in lanes 14-18, addition of BrdUTP and IodUTP to a '-C' reaction caused increased primer elongation by *E. coli* DNA polymerase I, as expected if these analogs can pair with G residues in the template during DNA synthesis. From this experiment (and others, not shown) it is seen that replacement of dCTP by IodUTP occurs more readily than replacement by BrdUTP. In addition, IodUTP was incorporated to a lesser degree in place of dGTP (compare lanes 8 and 12). Not shown is that both analogs were very efficiently incorporated in place of dTTP (which represents the normal base pairing specificity for these derivatives). The arrow to the right of the '-C' lanes marks a strong 'hotspot' for misincorporation of BrdUTP and IodUTP (and perhaps a normal dNTP) on the RA69 template-primer.

We have measured nucleotide turnover (dNTP→dNMP) during 'minus' reactions to determine whether this activity of the polymerase could deplete the supply of dNTP during 'minus' reactions. Experiments with DNA polymerase I of *E. coli* (data not shown) showed that the extent of dNTP→dNMP turnover in reactions carried out in the presence of only 3 of the 4 dNTPs was insignificant.

DISCUSSION

We have devised an electrophoretic assay to examine the fidelity of DNA synthesis in the region of a natural template near the 3'-OH of a discrete primer, annealed to the template. This experimental approach takes advantage of the fact that severe imbalance in dNTP concentrations greatly increases the error frequency during *in vitro* DNA synthesis (4,23-25). The technique that we have used to study misincorporation carries pool imbalance to the extreme (complete omission of one of the 4 dNTPs from the polymerase reaction). This ploy causes an extremely rare event (misincorporation) to occur sufficiently frequently to permit its biochemical detection. Because the rate of primer elongation is limited by the frequency of stable incorporation of noncomplementary nucleotides, each primer commonly undergoes several misincorporation events during the polymerase reaction. By carrying out all four 'minus' reactions in one experiment we can simultaneously compare the propensity for misincorporation at several sites on the template. The use of primers that anneal at different regions of the template allows this comparison to be extended to many positions along a natural template. Because of the severely nonphysiological condition imposed on the polymerization reaction (absence of one dNTP) we have not attempted to translate our data into quantitative estimates of error frequency. Nonphysiological reaction conditions are routinely used in many enzymological studies of catalytic capability and substrate specificity. We therefore regard our technique as a valid probe of an important property of DNA polymerases, their accuracy. The

electrophoretic assay represents a sensitive method for identification of reaction conditions that influence the accuracy of DNA synthesis. Most importantly, the approach described here provides the means for extensive characterization of the sequence dependence of mispairing during DNA synthesis. Finally, the experimental system described here permits detection of mispairing during DNA synthesis, potentiated by specific mutagen-induced base modifications.

Regarding the latter capability, we have used the electrophoretic assay of misincorporation to examine the base pairing properties of two commercially available dTTP analogs, BrdUTP and IodUTP. BrUra and BrUrd have been used extensively to produce mutations in procaryotic systems. It is widely believed that the molecular mechanism of BU mutagenesis involves mispairing between BU and G during DNA synthesis. This hypothesis arose from genetic studies, which characterized the base changes produced in DNA of cells treated with the analog (26-28), and from chemical considerations, which suggested that BU•G mispairing could be promoted by shifts in the tautomeric (26) or ionic (29) state of the analog. The data of Figure 4 support the hypothesis that BU•G mispairing can occur during DNA synthesis. Previous support for this type of mispairing came from early work of Trautner *et al.* (30), who obtained evidence that dGMP residues can be incorporated by *E. coli* pol I with a poly[dA-BrdU] template, but not with a poly[dA-dT] template. In studies conducted concurrently to ours (31), Lasken & Goodman have demonstrated conclusively that BU•G mispairing occurs during DNA synthesis catalyzed by purified DNA polymerase of bacteriophage T4, when BU is present on either incoming dNTP or template.

We believe that the experimental system described here constitutes a generally applicable approach for examining the molecular basis of mutagenesis induced by a wide variety of agents. Any chemically modified dNTP that is stable and highly purified can be tested for misincorporation at many different sites on a natural DNA template. In light of the recent finding (32) that DNA precursor pools are more susceptible to alkylation than are residues in DNA, the electrophoretic assay of misincorporation could be applied to any potentially mispairing dNTP derivative that would be formed *in vivo* by treatment of cells with a chemical mutagen. In addition, if a modified nucleotide can be incorporated *in vitro* into a growing primer, the newly synthesized strand can then be isolated and used as template to determine whether a given chemical modification imparts any ambiguous base pairing potential to the template.

The results shown in Figures 1, 2 and 4, combined with the data presented in the accompanying paper (Hillebrand & Beattie) clearly demonstrate the phenomenon of sequence-dependent variation in the accuracy of DNA synthesis. This phenomenon was previously demonstrated for *in vitro* replication catalyzed by the multienzyme replication complex of bacteriophage T4 (33). In that work the frequency of misincorporation at two different amber sites on the ϕ X174 genome was shown to differ, as measured by a transfection assay.

By use of the electrophoretic assay of misincorporation with *E. coli* DNA polymerase I, we have tentatively identified about twenty 'hotspots' for misincorporation. Although these 'hotspots' were assigned on the basis of highly qualitative data, several trends have nevertheless emerged with regard to features of the template near these sites. First, over half of the 'hotspots' for misincorporation by DNA polymerase I of *E. coli* were at T residues on the template (perhaps reflecting the phenomenon that overall, misincorporation

by this enzyme appears to occur most readily in place of dATP). The second emerging trend is the existence of a region of >70% [G+C] base composition on the immediate 5'-side of template residues displaying relatively high rates of misincorporation. A high [G+C] content in the region of DNA synthesized immediately following misincorporation may actually promote stable misincorporation with *E. coli* pol I, by stabilizing the mispair against subsequent removal by the 3'-exonuclease activity of the polymerase. Until the primer containing the mismatched terminus is elongated by several more residues, the misincorporated residue may be subject to this 'editing' activity; a high [G+C] content (imparting high stability) in the double helix formed immediately after misincorporation may serve to 'lock in' the replication error. Clearly, though, some other factor must be the major determinant of 'hotspots' for misincorporation, since high [G+C] content is absent from the 5'-side of many 'hotspots' and is present on the 5'-side in the case of many template positions displaying relatively low propensity for misincorporation.

Another apparent correlation between template sequence and propensity for misincorporation was revealed by experiments in which we used electrophoresis to monitor the rate of phosphodiester bond formation at different template positions in the presence of all 4 dNTPs (Hillebrand & Beattie, manuscript in preparation). In that work we found that the rate of incorporation of correct dNMP residues also varied a great deal at different template positions when primer elongation was catalyzed by *E. coli* DNA polymerase I. However, even from the limited amount of data available on elongation with 3 vs 4 dNTPs on the same template, it is clear that there is no strong correlation between the rate of correct incorporation and the rate of incorrect incorporation at different template positions. However, there is a strong tendency for 'hotspots' for misincorporation to exist immediately to the 3'-side of template positions at which phosphodiester bond formation in the presence of all 4 dNTPs occurs slowly. Thus, in general, at template positions exhibiting the greatest propensity for misincorporation, incorporation of the subsequent (correct) dNMP residue is inherently slow, although not slow enough to allow accumulation of mismatch-terminated intermediates that can be visualized by electrophoresis/autoradiography. Our present understanding of the DNA polymerase reaction mechanism does not permit a rational explanation for this finding.

Interestingly, the 'postmisincorporation pause sites' do not correlate well with template positions of slowest phosphodiester bond formation in the presence of all 4 dNTPs. The phenomenon of postmisincorporation pause may be related to a previously observed characteristic of chain elongation on natural DNA templates damaged with chemicals and radiation (34,35). Results of electrophoretic analyses of primer extension catalyzed by purified DNA polymerases indicated that chain elongation usually terminated or strongly paused after formation of phosphodiester bonds at template residues immediately preceding damaged, 'noninstructional' template residues. However, with certain DNA polymerases and at certain template positions, elongation terminated or strongly paused after incorporation of a dNMP opposite the damaged template residue. Both this phenomenon and the existence of 'postmisincorporation pause sites' probably involve a complex, sequence-dependent set of interactions between the template-primer, DNA polymerase and incoming dNTPs. The data of Figure 3 suggest that strong base stacking interactions at the primer terminus promote the 'postmisincorporation pause' phenomenon, possibly by stabilization of the terminal mismatch against exonucleolytic removal, which may be one requirement for accumulation of these elongation intermediates.

From a systematic effort to identify misincorporated dNMPs and to obtain more quantitative data on the propensity for misincorporation at different sites within a much more extensive set of template sequences, we may eventually identify features of the template sequence that govern the frequency and nature of misincorporation. Perhaps related to this question, it is known (36) that certain types of mispairs can exist within a normal B-DNA geometry, provided that the mispair is surrounded by the proper base pairs (presumably those providing sufficient base stacking interactions to overcome the energetically unfavorable tautomeric form of one member of the hydrogen-bonded mispair). A full explanation of the sequence dependence of misincorporation would also likely involve the DNA polymerase. In a misincorporation event, in order for an incorrect nucleotide to bind to the template-primer-polymerase complex in the correct orientation for phosphodiester bond formation to take place, stabilization energy may have to derive not only through base stacking interaction, which in this case could only occur on one side of the mispaired dNTP, but also through interactions with the DNA polymerase and divalent cation. An additional factor in the case of many DNA polymerases (*E. coli* pol I included) is the 3'-exonuclease 'proofreading' activity of the polymerase, which may also be modulated by complex, sequence-dependent interactions.

ACKNOWLEDGEMENTS

Work supported by Grants GM25530 and GM30590 from the National Institutes of Health, Grant Q819 from the Robert A. Welch Foundation and Grant 490 from the Anna Fuller Fund. G.G.H. was recipient of a Robert A. Welch Predoctoral Fellowship. K.L.B. is recipient of Research Career Development Award CA00891 from the National Cancer Institute. The authors thank Raymond Fenwick, Andrew Alpert and Charles Largent for advice and assistance in HPLC. A special acknowledgement is given to Dr. David Ward, who first suggested the use of the electrophoretic assay of chain elongation for investigating the incorporation of nucleotide analogs.

+ Present address: Department of Biochemistry, University of Illinois, Urbana, IL 61801, USA

*To whom correspondence should be addressed

REFERENCES

1. Hall, Z.W. and Lehman, I.R. (1968) *J. Mol. Biol.* **36**, 321-333.
2. Trautner, T.A., Swartz, M.N. and Kornberg, A. (1962) *Proc. Natl. Acad. Sci., U.S.A.* **48**, 449-455.
3. Weymouth, L.A. and Loeb, L.A. (1978) *Proc. Natl. Acad. Sci., U.S.A.* **75**, 1924-1928.
4. Kunkel, T.A. and Loeb, L.A. (1979) *J. Biol. Chem.* **254**, 5718-5725.
5. Bessman, M.J., Muzyczka, N., Goodman, M.F. and Schmaar, R.L. (1974) *J. Mol. Biol.* **88**, 409-421.
6. Goodman, M.F., Gore, W.C., Muzyczka, N. and Bessman, J. (1974) *J. J. Mol. Biol.* **88**, 423-435.
7. Goodman, M.F., Hopkins, R. and Gore, W.O. (1977) *Proc. Natl. Acad. Sci., U.S.A.* **74**, 4806-4810.
8. Loeb, L.A. and Kunkel, T.A. (1982) *Ann. Rev. Biochem.* **52**, 429-457.
9. Beattie, K.L., Wiegand, R.C. and Radding, C.M. (1977) *J. Mol. Biol.* **116**, 783-803.
10. Messing, J. (1983) *Meth. Enzymol.* **101**, 20-78.
11. Cunningham, R.P., DasGupta, C., Shibata, T. and Radding, C.M. (1980) *Cell* **20**, 223-235.
12. Godson, G.N., Barrell, B.G., Staden, R. and Fiddes, J.C. (1978) *Nature* **276**, 236-247.
13. van Wezenbeek, P.M.G.F., Hulsebos, T.J.M. and Schoenmakers, J.G.G. (1980) *Gene* **11**, 129-148.
14. Rowen, L. and Kornberg, A. (1978) *J. Biol. Chem.* **253**, 758-764.
15. Smith, M. and Khorana, H.G. (1963) *Meth. Enzymol.* **6**, 645-669.

16. Shortle, D. and Botstein, D. (1982) in **Molecular and Cellular Mechanisms of Mutagenesis**, Lemontt, J.F. and Generoso, W.M. Eds., pp. 147-156, Plenum Press, New York.
17. Shortle, D., Grisafi, P., Benkovic, S. and Botstein, D. (1982) *Proc. Natl. Acad. Sci., U.S.A.* **79**, 1588-1592.
18. Zakour, R.A. and Loeb L.A. (1982) *Nature* **295**, 708-710.
19. Brutlag, D. and Kornberg, A. (1972) *J. Biol. Chem.* **247**, 241-248.
20. Reckmann, B., Grosse, F. and Krauss, G. (1983) *Nucleic Acids Res.* **11**, 7251-7260.
21. Frank, R. and Koster, H. (1979) *Nucl. Acids Res.* **6**, 2069-2086.
22. Ornstein, R.L. and Fresco, J.R. (1983) *Biopolymers* **22**, 1979-2000.
23. Fersht, A.R. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4946-4950.
24. Kunkel, T.A. and Loeb, L.A. (1980) *J. Biol. Chem.* **255**, 9961-9966.
25. Kunkel, T.A., Silber, J.R. and Loeb, L.A. (1982) *Mutat. Res.* **94**, 413-419.
26. Freese, E. (1959) *J. Mol. Biol.* **1**, 87-105.
27. Champe, S.P. and Benzer, S. (1962) *Proc. Nat. Acad. Sci., USA* **48**, 532-546.
28. Drake, J.W. (1970) **The Molecular Basis of Mutation**, Holden-Day, San Francisco.
29. Lawley, P.D. and Brookes, P. (1962) *J. Mol. Biol.* **4**, 216-219.
30. Trautner, T.A., Swartz M.N. and Kornberg, A. (1962) *Proc. Natl. Acad. Sci., U.S.A.* **48**, 449-455.
31. Lasken, R.S. and Goodman, M.F. (1984) *J. Biol. Chem.*, in press.
32. Topal, M.D. and Baker, M.S. (1982) *Proc. Nat. Acad. Sci., USA* **79**, 2211-2215.
33. Hibner, U. and Alberts, B.M. (1980) *Nature* **285**, 300-305.
34. Moore, P.D., Bose, K.K., Rabkin, S.D. and Strauss, B.S. (1981) *Proc. Natl. Acad. Sci., U.S.A.* **78**, 110-114.
35. Strauss, B., Rabkin, S., Sagher, D. and Moore, P. (1982) *Biochimie* **64**, 829-838.
36. Fresco, J.R., Broitman, S. and Lane, A.-E. (1980) in **Mechanistic Studies of DNA Replication and Genetic Recombination**, Alberts, B. Ed., pp. 753-768, Academic Press, New York.