In Vitro Evidence That UV-Induced Frameshift and Substitution Mutations at T Tracts Are the Result of Misalignment-Mediated Replication Past a Specific Thymine Dimer†

Cheng-I. Wang and John-S. Taylor*
Department of Chemistry, Washington University, St. Louis, Missouri 63130
Received September 24, 1991; Revised Manuscript Received January 23, 1992

ABSTRACT: A previous study of UV-induced (254 nm) mutations in the *lacI* gene of *Escherichia coli* found that frameshift mutations accounted for about 35% of the observed mutations and that these mutations occurred predominantly at A_n·T_n sequences [Miller, J. H. (1985) *J. Mol. Biol. 182*, 48–65]. Because A_n·T_n sequences are hotspots for cis–syn thymine dimer formation [Brash, D. E., & Haseltine, W. A. (1982) *Nature 298*, 189–192], it would appear that UV-induced frameshift mutations are the result of an error during replicative bypass of a thymine dimer within such a sequence. To test the validity of such a proposal, replication experiments were carried out on templates containing cis–syn thymine dimers at each of the five possible sites of a T_6 tract. The 59-mer templates were prepared by ligating oligonucleotides containing an *Eco*RI site to the 5′-end of decamers containing the cis–syn thymine dimer and oligonucleotides containing the primer site to the 3′-end. Primer-extension reactions were then carried out on these templates with a 3′ → 5′ exonuclease-deficient (exo⁻) Klenow fragment of *E. coli* polymerase I and an exo⁻ T7 polymerase (Sequenase Version 2.0). The replicative bypass products were cleaved with *Eco*RI to rigorously establish and quantify the presence of frameshift mutations. Both polymerases were able to bypass dimers at all sites, but only the exo⁻ T7 polymerase led to detectable frameshifts, both −1 (~30%) and −2 (~5%), and only with the template containing a cyclobutane dimer at the second site from the 5′-end of the T_6 tract. Sequencing of the T7 polymerase-catalyzed bypass products of all templates demonstrated that within the limits of discrimination only As were introduced opposite the dimer-containing T tracts. The only exception was for the template with the dimer at the second site which led to a readily detectable amount of a substitution mutation (~30%) opposite the 5′-thymine of the T_6 tract. A mechanism involving a competition between reversible misalignment and realignment steps and irreversible elongation steps is proposed to explain the origin of both the frameshift and the substitution mutations. The implications of this work to the mechanism of UV-induced frameshift and substitution mutations at T tracts in vivo are discussed.

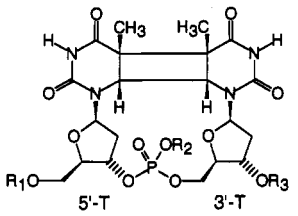
The correlation between sunlight and skin cancer has been well established and is thought to arise from mutations in protooncogenes induced by the UV[1] wavelengths present in sunlight [for a review, see Ananthaswamy and Pierceall (1990)]. Ultraviolet radiation is known to induce a wide variety of substitution, frameshift, and deletion mutations in both prokaryotic [for a review, see Hutchinson (1987)] and eukaryotic systems [see, for example, Hsia et al. (1989)]. Most of the substitution mutations occur at dipyrimidine sequences, as do the frameshift mutations, which occur primarily at homopyrimidine sequences. The major photoproducts of duplex DNA induced by UV radiation also occur at dipyrimidine sites and are the cis–syn cyclobutane dimer, the (6–4) pyrimidine–pyrimidinone product, and its valence isomer, the Dewar photoproduct [for a review, see Cadet and Vigny (1990)]. That many of the sites of UV-induced mutations and DNA photoproducts coincide has led to the conclusion that most UV-induced mutations are the result of repair and/or replication errors made at the site of DNA photoproducts (targeted mutations). Mutation hotspots do not correlate with photoproduct formation hotspots, however, suggesting that the frequency and type of errors made are sequence-dependent (Brash et al., 1987). The mechanisms by which these errors occur, and how and why they depend on sequence, are poorly understood, primarily because of the lack of general methods for preparing pure, well-characterized, site-specific photoproduct-containing DNA. We are approaching this problem by utilizing synthetic, enzymological, and molecular biological methods to construct site-specific photoproduct-containing DNA for precise physical and biological studies designed to address these issues [for a general description of our approach, see Taylor (1990)].

One major class of mutations induced by UV radiation are frameshift mutations. T tracts in the *lacI* gene of *Escherichia coli* are hotspots for cis–syn thymine dimer (**1a**; Figure 1) formation (Brash & Haseltine, 1982) and −1 frameshift mutations (Miller, 1985; Schaaper et al., 1987). Together, these facts suggest that the −1 frameshift mutations are the result of an error made during replication of a T tract containing a cis–syn thymine dimer. It is also possible, however, that some or all of the frameshift mutations are due to nontargeted mutations caused by error-prone replication induced by the SOS response (Wood & Hutchinson, 1984). A strand slippage mechanism was originally proposed by Streisinger and coworkers to account for spontaneous −1 frameshift mutations

† This investigation was supported by PHS Grant No. RO1-CA40463, awarded by the National Cancer Institute, DHHS.
* To whom correspondence should be addressed.

[1] Abbreviations: ATP, adenosine triphosphate; BME, β-mercaptoethanol; BSA, bovine serum albumin; DMT, 4′,4′-dimethoxytrityl; EDTA, ethylenediaminetetraacetic acid; ddNTP, 2′,3′-dideoxynucleotide triphosphate; dNTP, 3′-deoxynucleotide triphosphate; DTT, dithiothreitol; exo⁻, 3′ → 5′ exonuclease deficient; KF, Klenow fragment of *E. coli* polymerase I; NMR, nuclear magnetic resonance; UV, ultraviolet radiation.

0006-2960/92/0431-3671$03.00/0



1a R_1=5'-end, R_2=(-), R_3=3'-end

1b R_1=DMT, R_2=CH_3, R_3=P(OCH_3)(N(i-Pr)$_2$)

FIGURE 1: The structure of the cis–syn thymine dimer (**1a**) and the building block used to incorporate it site-specifically into DNA (**1b**).

at undamaged homopolymer tracts (Streisinger et al., 1966) and was later elaborated to account for a wide variety of spontaneous frameshift and substitution mutations [for recent reviews, see Kunkel (1990) and Ripley (1990)]. Strand slippage was also suggested to be a mechanism whereby a dimer could cause a frameshift to occur at homopolymer tracts (Schaaper et al., 1987). It has also been proposed that UV-induced frameshift mutations might be the result of the replicative bypass of a cyclobutane dimer formed between nonadjacent thymines (Nguyen & Minton, 1988). Unfortunately, it is impossible to deduce from mutation spectra alone which, if any, of the $n-1$ possible adjacent or the $n-2$ nonadjacent dipyrimidine photoproducts of a homopyrimidine tract of length n are responsible for the observed mutations.

Our approach to unraveling the origin of UV-induced frameshift mutations is to determine the structure and properties of T tracts containing site-specific photoproducts and to study their repair and replication both in vitro and in vivo. T tracts, otherwise known as A tracts, are particularly interesting because they appear to have a unique structure which causes heterogeneous DNA containing them to bend [for recent reviews, see Crothers et al. (1990) and Hagerman (1990)]. We recently determined the site-specific effect of dimer formation on A tract bending by synthesizing all possible cis–syn thymine dimer monoadducts of a T_6 tract-containing decamer (Figure 2) utilizing the cis–syn thymine dimer building block **1b** (Figure 1). Herein, we report the use of the same dimer-containing decamers used in the bending study to construct 59-mer templates containing all five possible thymine dimer monoadducts of a T_6 tract. In vitro replication of these templates with two exo-deficient (exo$^-$) polymerases led to dimer-site- and polymerase-specific production of both −1 and −2 frameshift mutations and a unique substitution mutation. A mechanism involving reversible misalignment and realignment steps in competition with irreversible polymerization steps is proposed to account for the formation of the observed mutations in vitro, and it may explain the origin of the corresponding UV-induced mutations in vivo.

Materials and Methods

Enzymes and Reagents. T4 polynucleotide kinase was from Bethesda Research Labs. T4 DNA ligase was from New England Biolabs. Sequenase Version 1.0 (Tabor & Richardson, 1987) and Version 2.0 [Δ28(Lys^{118}–Arg^{145}) (Tabor & Richardson, 1989)] and 3′ → 5′ exo$^-$ KF [D355A E357A (Derbyshire et al., 1988)] were from USB. *Eco*RI was from Promega; dNTPs and ddNTPs were from Fisher, and [γ-^{32}P]ATP was from Amersham. *E. coli* photolyase was obtained from A. Sancar (University of North Carolina), and T4 *denV* endonuclease V was from S. Lloyd (Vanderbilt University).

Construction of the Templates. Oligonucleotides were synthesized by automated DNA synthesis and purified either by reverse-phase HPLC or by polyacrylamide gel electrophoresis. Dimer-containing oligonucleotides were previously synthesized for use in another study (Wang & Taylor, 1991) with a thymine dimer building block (Taylor et al., 1987) (structure **1b**, Figure 1). The dimer-containing decamer and the appropriate 3′-oligonucleotide (2 nmol of each) (Figure 2) were 5′-phosphorylated with 10 units of T4 polynucleotide kinase at 37 °C for 30 min in 25 μL of 70 mM Tris-HCl (pH 7.6)/10 mM $MgCl_2$/5 mM DTT/2 mM ATP. The two phosphorylated oligonucleotides and 2 nmol of the appropriate

Replication Template Subunits:

24-mer	5′-AGCTACCATGCCTGCATTGAATTC-3′
T_6	5′-pTGCCGTTTTT-3′
DT_4	5′-pTGCCGT=TTTT-3′
TDT_3	5′-pTGCCGTT=TTT-3′
25-mer	5′-pTGCCGTCGTAATCATGGTCATAGCT-3′
26-mer	5′-AGCTACCATGCCTGCATTGAATTCTG-3′
T_2DT_2	5′-pCCGTTT=TTTG-3′
T_3DT	5′-pCCGTTTT=TTG-3′
23-mer	5′-pCCGTCGTAATCATGGTCATAGCT-3′
30-mer	5′-AGCTACCATGCCTGCATTGAATTCTGCCGT-3′
T_4D	5′-pTTTT=TGCCGT-3′
19-mer	5′-pCGTAATCATGGTCATAGCT-3′

Ligation Scaffold:

40-mer	3′-CGGACGTAACTTAAGACGGCAAAAAACGGCAGCATTAGTA-5′

FIGURE 2: The oligonucleotides used to construct the site-specific cis–syn thymine dimer-containing templates used in the replicative bypass experiments. The 40-mer ligation scaffold was used to hold the 5′-oligonucleotide, the dimer-containing decamer, and the 3′-flanking oligonucleotide of each set in place during ligation with T4 DNA ligase and ATP.

5′-oligonucleotide were annealed to 1.5 nmol of the 40-mer ligation scaffold by heating the mixture to 90 °C in a 300-mL water bath which was then allowed to slowly cool to 16 °C. A total of 800 units of T4 DNA ligase was then added followed by ATP to a final concentration of 2.5 mM. The ligation reactions were allowed to proceed at 16 °C overnight and were then electrophoresed on a 15% preparative denaturing polyacrylamide gel. The desired ligation products (59-mers) were excised from the gel, electroeluted, and dialyzed.

Characterization of the Dimer-Containing Templates. T4 *denV* endonuclease V cleavage: Approximately 0.5 pmol of each dimer-containing template was 5′-end labeled and annealed to 0.8 pmol of unlabeled complementary 40-mer. A total of 0.1 pmol of each double-stranded substrate was incubated with 112 units of T4 *denV* endonuclease V in 10 μL of 32 mM Tris-HCl (pH 8.3)/10 mM EDTA/100 mM NaCl/0.1 mg/mL BSA, at 30 °C. After 40 min, each reaction was added to 20 μL of 1.5 M aqueous piperidine and the mixture was heated at 95 °C for 25 min. The samples were lyophilized, precipitated from 1 M LiCl with 6 volumes of 1:3 ethanol/acetone, and electrophoresed on a 15% polyacrylamide gel.

Sequencing of photoreversed TDT_3 template: 10 pmol of the TDT_3 template was photoreversed by incubating with 3 pmol of *E. coli* photolyase in 40 μL of 100 mM NaCl/50 mM Tris-HCl (pH 7.4)/1 mM EDTA/20 mM dithiothreitol/0.1 mg/mL BSA/4% glycerol in an ice bath at a distance of approximately 5 cm from a 135-W incandescent bulb. The reaction mixture was phenol extracted and precipitated with 1 M LiCl and 6 volumes of 1:3 ethanol/acetone. Dideoxy sequencing (Figure 3) was carried out on the photoreversed TDT_3 and T_6 templates as described below with 1 pmol of 5′-labeled 15a-mer annealed to 1.5 pmol of template and ddNTP:dNTP ratios of 1:5, 1:3, 1:3, and 1:1 for A, C, G, and T respectively.

Replication Experiments. Primer-templates were prepared by annealing 0.5 pmol of 5′-labeled 15a-mer primer to 1 pmol of each unlabeled template. Primer-extension reactions catalyzed by Sequenase Version 2.0 were conducted at 22, 37, or 52 °C for 2.5 h with 2.5 units of enzyme in 15 μL of 40 mM Tris-HCl (pH 7.5)/20 mM $MgCl_2$/50 mM NaCl/7 mM DTT/200 μM dNTPs. Primer-extension reactions catalyzed by exo$^-$ KF were conducted at 37 °C for 2.5 h with 4 units of enzyme in 15 μL of 10 mM Tris-HCl (pH 7.4)/10 mM $MgCl_2$/50 mM NaCl/10 mM BME/200 μM dNTPs. The reaction mixtures were then phenol extracted twice, lyophilized, and precipitated from 1 M lithium chloride with 6 volumes of 1:3 ethanol/acetone. A total of 0.1 pmol of the ethanol-precipitated DNA from each primer-extension reaction was treated with 2 units of *Eco*RI at 37 °C for 1.5 h in 10 μL of 90 mM Tris-HCl (pH 7.5)/10 mM $MgCl_2$/50 mM NaCl/0.1 mg/mL BSA. The reaction mixtures were then lyophilized, precipitated in the same way as described above, and electrophoresed on a 15% denaturing polyacrylamide gel. Product distributions were quantified by densitometry of autoradiograms obtained without the use of intensifier screens. Initial bypass rates were determined by linear regression analysis of the natural log of the fraction of termination products versus time for measurements made at 4, 8, 16, 24, and 32 min.

Isolation and Sequencing of the Bypass Products. Primer-extension reactions were conducted as described above but on a larger scale. Thus, 11 pmol of the 15a-mer primer was annealed to 10 pmol of each template and incubated at 37 °C for 2.5 h with 45 units of Sequenase Version 2.0 in a total volume of 300 μL. The reaction mixtures were then phenol extracted, made up to 1 M LiCl, precipitated with 6 volumes of 1:3 ethanol/acetone, and electrophoresed on an analytical 15% acrylamide denaturing gel. The gel section corresponding to the bypass product of each reaction was excised and electroeluted using an electrophoretic concentrator (ISCO Model 1750). The eluates were then dialyzed against water overnight. About one-tenth of each bypass product was subjected to primer-extension with 1 unit of Sequenase Version 1.0 and 1 pmol of 5′-^{32}P-end-labeled 15c-mer at 37 °C for 15 min in 10 μL of 40 mM Tris-HCl (pH 7.5)/20 mM $MgCl_2$/50 mM NaCl/7 mM DTT/50 μM dTTP. The reactions were quenched by adding EDTA to a final concentration of 15 mM. One-fifth of each bypass product was sequenced by the dideoxy termination method with 2 units of Sequenase Version 1.0 and 1 pmol of 5′-^{32}P-end-labeled 24-mer in 10 μL of 40 mM Tris-HCl (pH 7.5)/20 mM $MgCl_2$/50 mM NaCl/7 mM DTT/250 μM ddNTP/dNTP mixture. The molar ratios of ddNTP to dNTP were 1:1, 1:3, 1:3, and 1:5 for A, C, G, and T, respectively. The reactions were incubated at 37 °C for 15 min and were quenched by adding EDTA to a final concentration of 15 mM.

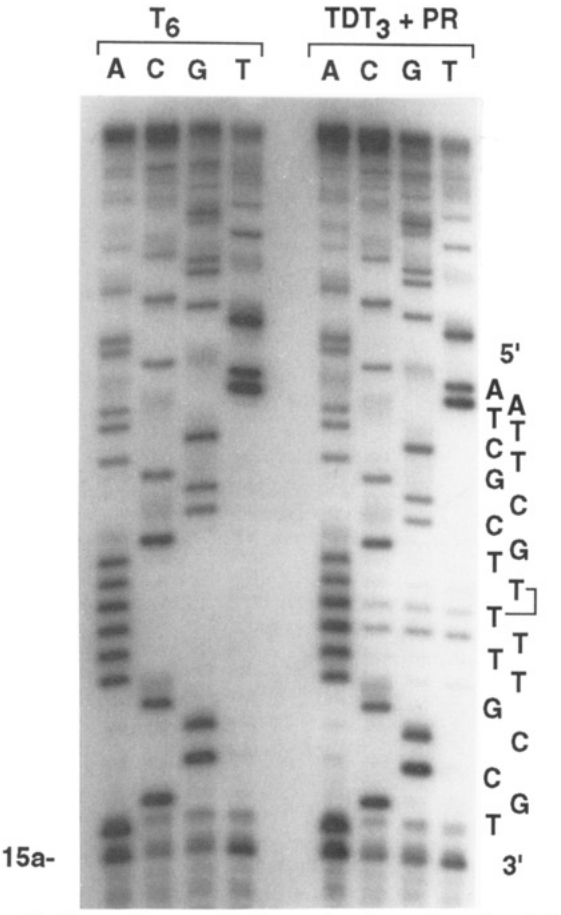


FIGURE 3: Autoradiogram of a denaturing electrophoresis gel of the dideoxy sequencing products for the T_6 and the photoreversed (PR) TDT_3 59-mer templates. The lanes are labeled according to the dideoxynucleotide used in the sequencing reactions. The sequence to the right is that of the 59-mer templates with the position of the thymine dimer indicated by a bracket. The presence of termination bands in the TDT_3 sequencing lanes corresponding to the T_6 tract are the result of a small amount (~5%) of non-photoreactivated dimer. Virtually identical distributions of smeared bands are observed in the sequencing lanes of both templates and are attributed to incomplete denaturation of secondary structures.

RESULTS

Design, Synthesis, and Characterization of the Templates. The replicative bypass of DNA damage can best be studied

Table I: Percentage Distribution of Elongated Products Resulting from the Replication of Dimer-Containing Templates by Sequenase Version 2.0 and by exo⁻ KF[a]

product	T (°C)	T_6	DT_4	TDT_3	T_2DT_2	T_3DT	T_4D
bypass	22	>98	65.9 ± 4.5	72.0 ± 3.0	75.0 ± 1.8	28.8 ± 1.7	79.9 ± 3.0
	37	>98	57.6 ± 0.8	45.5 ± 4.7	66.9 ± 1.6	29.7 ± 2.3	25.1 ± 2.5
	52	>98	20.2 ± 0.3	20.8 ± 1.0	38.2 ± 1.7	9.3 ± 0.8	5.2 ± 0.4
	37	(>98)	(39.3 ± 0.2)	(65.7 ± 3.4)	(73.4 ± 1.7)	(69.7 ± 0.8)	(36.2 ± 4.0)
termination	22	<1	<1	5.9 ± 0.4	6.0 ± 1.0	5.7 ± 0.3	10.2 ± 1.1
opp 5′-T	37	<1	<1	6.6 ± 1.4	3.1 ± 0.4	<1	5.9 ± 1.0
	52	<1	<1	6.8 ± 0.2	7.2 ± 0.8	<1	3.8 ± 1.6
	37	(<1)	(<1)	(5.5 ± 0.7)	(3.0 ± 0.4)	(<1)	(22.0 ± 4.3)
termination	22	<1	26.4 ± 2.8	19.7 ± 2.2	14.6 ± 0.3	37.0 ± 0.5	16.9 ± 1.0
opp 3′-T	37	<1	24.7 ± 0.2	31.6 ± 2.4	18.9 ± 1.2	19.9 ± 0.8	50.9 ± 4.7
	52	<1	48.0 ± 0.4	58.6 ± 1.5	38.4 ± 1.6	16.8 ± 1.7	39.3 ± 0.6
	37	(<1)	(55.7 ± 1.7)	(28.8 ± 3.8)	(17.7 ± 0.8)	(30.3 ± 0.8)	(41.8 ± 1.7)
termination	22	<1	7.7 ± 1.7	2.4 ± 0.7	4.4 ± 0.8	28.5 ± 1.7	3.0 ± 0.9
1 prior 3′-T	37	<1	17.7 ± 1.1	16.3 ± 1.7	11.1 ± 0.3	50.4 ± 2.2	18.1 ± 1.4
	52	<1	31.8 ± 0.6	13.8 ± 1.0	16.2 ± 0.9	73.9 ± 1.1	51.7 ± 0.9
	37	(<1)	(5.0 ± 1.6)	(<1)	(5.9 ± 1.7)	(<1)	(<1)

[a] Data for exo⁻ KF are given in parentheses. Reactions were conducted as described in the experimental. Averages of three experiments are given along with their standard deviations.

in vitro by utilizing site-specifically damaged templates of sufficient length to enable primer-extension reactions to be carried out on the template and dideoxy sequencing reactions to be carried out on the bypass products (Taylor & O'Day, 1990). Because 3′ → 5′ exonuclease-deficient polymerases often catalyze the addition of one or more nucleotides to the blunt end of DNA duplexes (Clark et al., 1987), we decided to introduce a restriction site into the templates on the 5′-side of the DNA damage site so that any bypass products could be made to have uniform 3′-ends (Figure 4). With these basic design criteria in mind, we chose to construct 59-mer templates with the dimer-containing T_6 tracts at the center and an *Eco*RI restriction site 10 nucleotides from the 5′-end of the T tract. The length of the template and positioning of the dimer were chosen so that primers 15 nucleotides in length could be used to prime both the extension and sequencing reactions at a sufficient distance from the dimer site to facilitate interpretation of the electrophoresis gel data.

The 59-mer templates were constructed by ligating the appropriate oligonucleotides to the 5′- and 3′-ends of dimer-containing decamers (Wang & Taylor, 1991) with T4 DNA ligase in the presence of ATP and a ligation scaffold (Figure 2). The site and integrity of the dimer in each template was confirmed by treatment with the pyrimidine dimer-specific T4 *denV* endonuclease V, which resulted in the expected cleavage products in 97.8%, 97.6%, 97.1%, 97.8%, and 97.7% yield for DT_4, TDT_3, T_2DT_2, T_3DT, and T_4D respectively. Because an A → C mutation at the 3′-end of the A tract (~30%) was later observed in the replication of the TDT_3 template, this template was also dideoxy sequenced following photoreversal along with the T_6 template (Figure 3). With the exception of termination bands due to the presence of a small amount of non-photoreactivated dimer (<5%), the autoradiograms for both templates are virtually identical and do not indicate the presence of a G at the 5′-end of the T tract. This result, together with the fact that replication of the T_6 template did not lead to any discernible substitution mutation (Figure 7), establishes the structural integrity of the TDT_3 template.

Replication Experiments. Primer-extension reactions were carried out on the T_6 tract containing templates with the 15a-mer primer according to the scheme in Figure 4. In addition to termination bands, which correspond to the arrest of chain elongation prior to, or opposite the 3′- or 5′-thymine of the dimer, bands corresponding to *Eco*RI-digested bypass products were observed (Figure 5). The relative amounts of bypass and termination products were found to depend on the dimer site, temperature, and polymerase used (Table I). The initial rate of dimer bypass catalyzed by Sequenase Version 2.0 at 37 °C was found to follow first-order kinetics with rate constants (h^{-1}) of 2.76 ± 0.15 for DT_4, 0.54 ± 0.04 for TDT_3, 1.26 ± 0.05 for T_2DT_2, 1.14 ± 0.09 for T_3DT, and 0.90 ± 0.04 for T_4D. Bypass of the dimers in DT_4, T_2DT_2, T_3DT, and T_4D (lanes 1, 3, 4, and 5 of Figure 5) by Sequenase Version 2.0 did not result in any readily discernible (>1%) frameshift mutations as indicated by the formation of single bands corresponding to 32-mers. On the other hand, bypass of the dimer in TDT_3 resulted in both −1 and −2 frameshift mutations in addition to the full-length bypass product, as evidenced by multiple bands corresponding to 32-, 31-, and 30-mers (Figure 5 and Table II). The ratio of bypass to termination products decreased as reaction temperatures rose from 22 to 52 °C. No readily discernible termination or frameshift products were observed in the replication of the dimer-free template in the same temperature range; nor were any readily discernible frameshift mutations detected at 37 °C for any of the templates when replicated with 3′ → 5′ exo⁻ KF, again indicating that the frameshift mutations observed for the TDT_3 template were a function of the polymerase and not due to synthesis errors in the TDT_3 template.

Table II: Percentage Distribution of Bypass Products of TDT_3 Produced by Sequenase Version 2.0[a]

	primer-extension/*Eco*RI cleavage procedure			TTP sequencing procedure
products	22 °C	37 °C	52 °C	37 °C
full length	54.5 ± 3.3	64.1 ± 1.2	59.1 ± 1.2	71.9 ± 2.7[b]
−1	32.4 ± 1.4	30.3 ± 1.3	36.8 ± 1.1	21.3 ± 0.5
−2	13.1 ± 1.8	5.6 ± 1.9	4.1 ± 0.8	6.7 ± 1.2

[a] Reactions were conducted as described in Materials and Methods. Averages of three experiments are given along with their standard deviations. [b] This value represents the sum of the percentage of +6 extension product (42.9 ± 2.7) and that of the unextended 15-mer primer (29.0 ± 1.4) in lane 2 of Figure 6. The latter product represents the percentage of the A → C substitution mutation.

Sequencing the Bypass Products. Information on the fidelity of the replication of the dimer-containing templates catalyzed by Sequenase Version 2.0 was first obtained by carrying out primer-extension reactions on each bypass product with the 15c-mer primer (Figure 4), dTTP, and Sequenase Version 1.0. Because Sequenase Version 1.0 has residual 3′ → 5′ exonuclease or proofreading activity (Tabor & Richardson, 1987) and T is the only nucleotide that can be incorporated, the number of sequential As that were incorporated

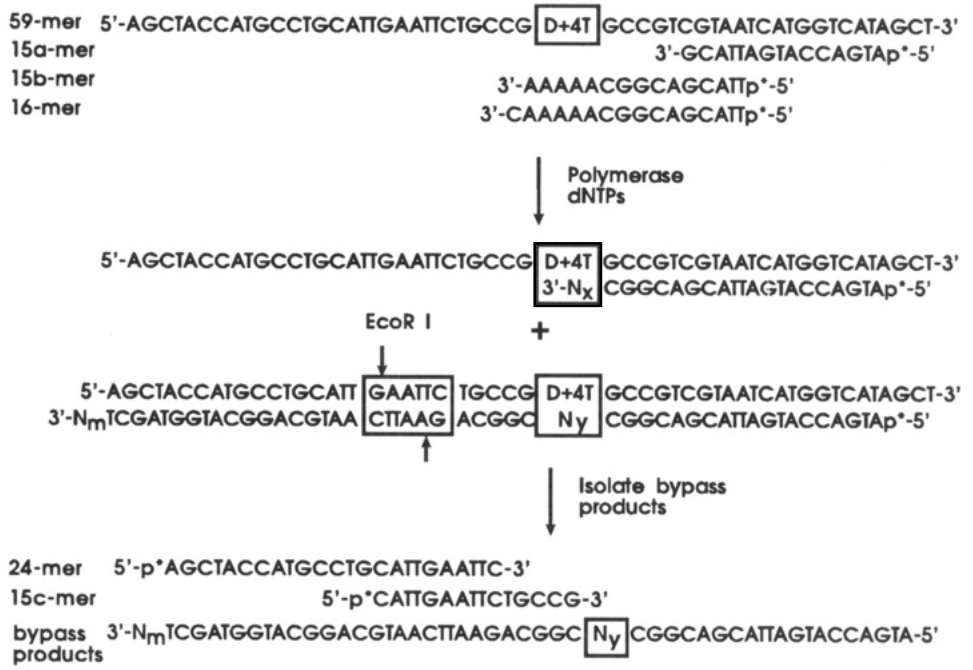


FIGURE 4: The steps and intermediates involved in assaying for frameshift and substitution mutations resulting from the replicative bypass of site-specific cis–syn thymine dimer-containing templates. D refers to a cis–syn thymine dimer, and N_x refers to the number (x) of nucleotides (N) inserted opposite the dimer-containing T tract. In the first assay, the products resulting from primer-extension opposite the 59-mer template were treated with *Eco*RI and subjected to denaturing gel electrophoresis (Figure 5). For the second and third assays, the bypass strand was isolated by denaturing gel electrophoresis. In the second assay, primer-extension was carried out with 15c-mer in the presence of dTTP only (Figure 6), and for the third assay standard dideoxy sequencing was carried out with the 24-mer primer (Figure 7).

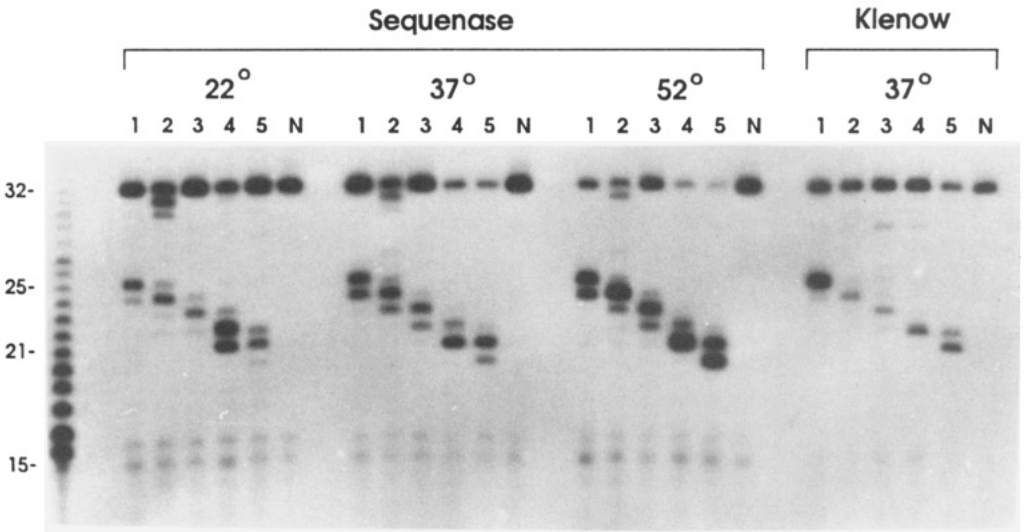


FIGURE 5: Autoradiogram of a denaturing electrophoresis gel of the products from the polymerase-catalyzed elongation of 5′-end labeled 15a-mer primer opposite the non-dimer and dimer-containing T tract templates followed by *Eco*RI treatment (Figure 4). 1, 2, 3, 4, 5, and N refer to the 59-mer templates DT_4, TDT_3, T_2DT_2, T_3DT, T_4D, and T_6. Reactions were carried out as described in Materials and Methods.

opposite the dimer-containing T_6 tracts in the bypass reaction can be readily determined. Only fully extended products were observed in significant amounts (we define significant to be >1% and estimate the detection limit to be ~1%), when primer-extension was carried out on the bypass products of T_6, DT_4, T_2DT_2, T_3DT, and T_4D (Figure 6), corresponding to the essentially error-free incorporation of six As opposite these non-dimer and dimer-containing T tracts. Had any significant amounts of substitution or frameshift mutations occurred, truncation products would have been observed. When primer-extension was carried out on the bypass products of TDT_3, the fully extended product and two truncated products were observed, corresponding to the incorporation of four, five, and six As opposite TDT_3. In addition to the formation of truncation products, it was observed that a sizable amount (29 ± 1.4%) of the 15c-mer primer had not undergone elongation. This could later be attributed to an A → C substitution mutation at the 3′-end of the A tract in the bypass product (vide infra). When corrected for the amount of the substitution mutation, the ratio of the three extension products was similar to that observed in the primer-extension reactions opposite TDT_3 (Table II), suggesting that the truncation products were primarily due to the presence of deletion rather than substitution mutations in the bypass product. Furthermore, the truncated products terminated at sites in the bypass product that do not correspond to the dimer site in the template and hence are unlikely to have been caused by substitution mutations.

To obtain additional sequence information on the fidelity of the replication reactions, the bypass products were subjected

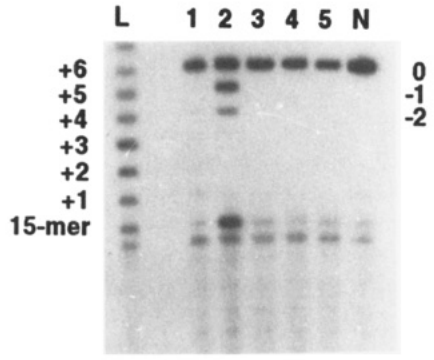


FIGURE 6: Autoradiogram of a denaturing electrophoresis gel of the products from elongation of the 15c-mer primer with dTTP opposite the replicative bypass products of the dimer and non-dimer-containing templates (Figure 4). 1, 2, 3, 4, 5, and N refer to the 59-mer templates DT_4, TDT_3, T_2DT_2, T_3DT, T_4D, and T_6. Numbers to the left correspond to the number of Ts added to the primer and those to the right correspond to the extent of frameshifting.

to dideoxy sequencing with Sequenase Version 1.0 and 5′-end-labeled 24-mer (Figure 4). The results (Figure 7) confirm that replication of DT_4, T_2DT_2, T_3DT, T_4D, and T_6 proceeded without readily discernible errors. Because the more sensitive dTTP sequencing experiment described above indicated the absence of significant (>1%) frameshift or substitution mutations, no attempt was made to quantify differences in background bands which were at a level of ~5%. Multiple bands were observed in sequencing the bypass products of TDT_3 indicative of −1 and −2 frameshifts. This can best be seen by comparing the dideoxy A lanes for the DT_4 and TDT_3 bypass products. At the top of the gel, two almost equally strong bands corresponding to a TT site in the template can readily be seen in the dideoxy A lane for DT_4 whereas four bands appear for TDT_3. The second from the top of these latter four appears approximately twice as dark as the first and third bands and much darker than the fourth band. This is what would be expected if the sequencing reactions for bypass products containing 0, −1, and −2 frameshift mutations were combined according to their frequency of occurrence. In addition, an extra band was observed in the dideoxy G lane, corresponding to an A → C transversion mutation at the 3′-end of the A_6 tract. This band (a 30-mer) was found to have 33% of the intensity of the neighboring band in the same lane (the 29-mer), compared to an average of 6 ± 1% for these same two bands in all the other sequenced templates. The amount of this extra G band compares quite favorably to that of the unelongated primer remaining in the primer-extension experiment with dTTP opposite this same sequence (29 ± 1.4%).

Discussion

We have discovered that replication of dimer-containing T_6 tracts in vitro can lead to significant amounts of frameshift and substitution mutations and that these mutations occur in a dimer-site- and polymerase-specific manner. Specifically, we have found that replication of TDT_3 by Sequenase Version 2.0 resulted in both −1 (~30%) and −2 (~5%) frameshift mutations and a single base substitution mutation (~30%), whereas replication of all of the other templates proceeded in a comparatively error-free manner. We have also found that no readily discernible frameshift mutations result when the same templates are replicated by exo⁻ KF. Evidence for the formation of frameshift and substitution mutations rested on gel electrophoretic analysis of (1) the products of primer-extension reactions opposite the dimer-containing templates following *Eco*RI cleavage, (2) the products of primer-extension reactions with dTTP opposite the A tract of the bypass products, and (3) the products resulting from dideoxy sequencing of the bypass products.

General Mechanistic Considerations. The significant amounts of −1 and −2 frameshift mutations and the single substitution mutation that occurred during replication of TDT_3 and their comparative absence in the replication of the other templates can be readily explained by a series of irreversible elongation steps (E) in competition with reversible misalignment (M) and realignment (R) steps (Figure 8). One key factor in our proposed mechanism is that the cis–syn thymine dimer is an instructive lesion that primarily codes for the addition of As opposite both the 5′- and 3′-Ts of the dimer. This is the case for thymine dimers flanked by purines when replicated in vitro by either pol I of *E. coli* or KF (Taylor & O'Day, 1990) and in vivo in an *E. coli* host (Banerjee et al., 1988). In accord with these results, both NMR and melting temperature studies indicate that the Watson–Crick hydrogen bonding properties of thymine are retained in the photo-

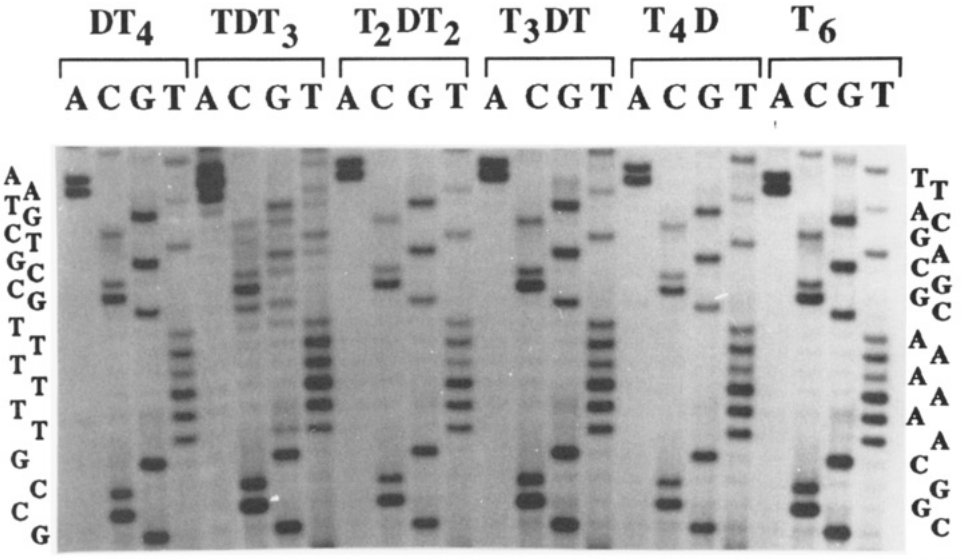


FIGURE 7: Autoradiogram of a denaturing electrophoresis gel of the products of Sanger dideoxy sequencing of the replicative bypass products of the non-dimer and dimer-containing T_6 tract templates using the 24-mer primer (Figure 4). The bases shown in the lane headings refer to the dideoxynucleotide used. The sequence written to the left is that of the T_6 template, and the sequence written to the right corresponds to that of the error-free replicative bypass of the T_6 template.

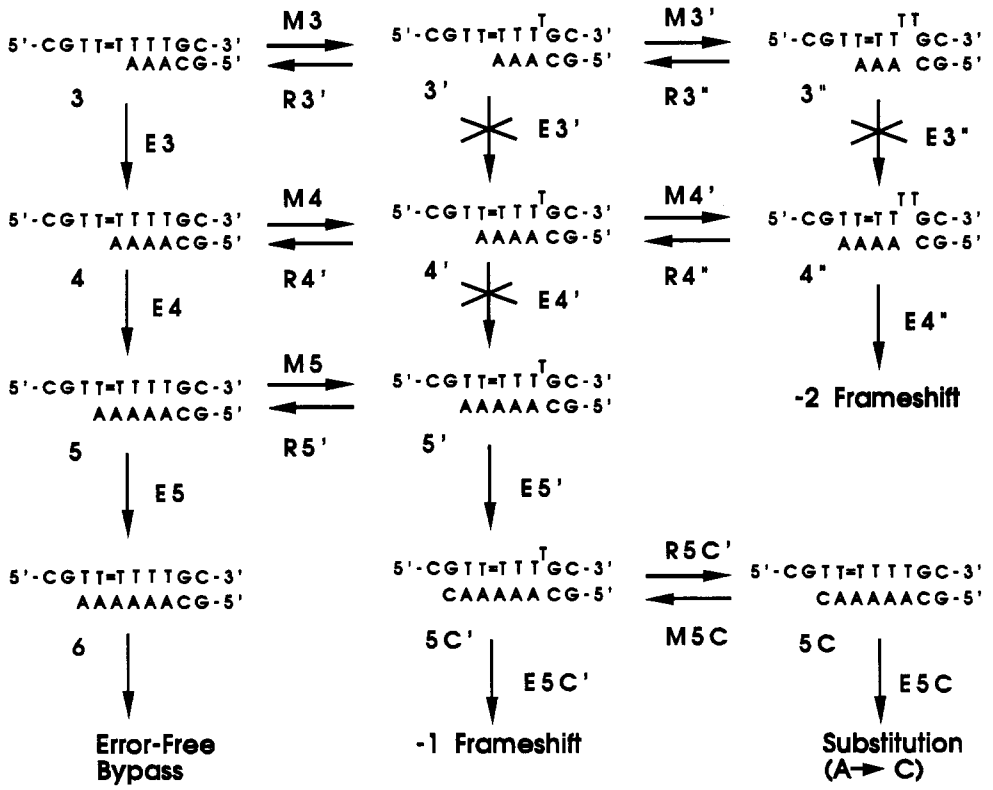


FIGURE 8: Proposed mechanistic scheme for the origin of the −1 and −2 frameshift and the A → C transversion mutations. Intermediates are labeled according to the number (n) of nucleotides in the A_n tract of the primer and according to the number of extra nucleotides in the template (denoted by the number of apostrophes). An intermediate in which the A_n tract of the primer is followed by another nucleotide (N) is denoted by nN. Vertical arrows represent irreversible chain elongation steps (Ei) catalyzed by the exonuclease-deficient polymerase, where i is the label of the intermediate being elongated. The horizontal arrows represent the reversible structural isomerization steps, misalignment (Mi) and realignment (Ri), where i is the label of the intermediate undergoing isomerization. Crossed-out steps indicate those steps that are not competitive with the other steps emanating from the same intermediate.

dimerization product (Kemmink et al., 1987; Taylor et al., 1990).

Another key factor in our mechanism is that elongation of primers terminating in A opposite both the 5′- or 3′-Ts of the dimer occurs at a rate much slower than for undamaged DNA (Taylor & O'Day, 1990). This results in the formation of primer-elongation products terminating one prior to, and opposite, the 3′-T of the dimer. The general absence of products terminating opposite the 5′-T of the dimer suggests that their elongation must occur at a rate more rapid than that of their formation. Preliminary results from our laboratories (Jiang and Taylor) indicate that for KF the rate of elongation of the primer terminating in A opposite the 5′-T of the dimer is approximately 40 times faster than that of the primer terminating in A opposite the 3′-T of the dimer. The latter elongation rate is approximately 3 orders of magnitude slower than for normal DNA and about 10 times slower than elongation of the primer terminating one prior to the dimer. The attenuation of the elongation rates by a dimer makes reversible structural rearrangements, such as misalignment, competitive with elongation of the aligned structure. Competitive misalignment caused by nucleotide misincorporation followed by elongation has been suggested to be the basis for spontaneous frameshift mutations (Bebenek & Kunkel, 1990). In our case, the formation of a Watson–Crick base pair between an A and the T of the dimer at a primer-terminus fulfills the same role as a mismatched base pair in the mechanism of spontaneous frameshift mutations.

Mechanism for the Formation of the −2 Frameshift Mutation. Given the general mechanistic considerations outlined above, we can propose a specific mechanism for the formation of the frameshift and substitution mutations in the replication of TDT_3 (Figure 8). In the first critical step, the primer terminating in an A_3 tract (3) is preferentially elongated by A opposite the 3′-T of the dimer (step E3) resulting in a primer-template (4) containing an A_4 tract properly aligned opposite the dimer-containing T tract. This intermediate can now follow either of two paths: further elongation (E4) or misalignment (M4). First, we will consider the consequences of the misalignment pathway, M4. Because the nucleotide immediately to the 5′-side of the dimer is a T, and both the 5′- and 3′-Ts of the dimer retain the base-pairing properties of T, two misaligned structures can form. The first misaligned structure (4′) corresponds to a one-base-pair slip, which forms a primer-terminus in which the 3′-end of the A_4 tract aligns with the 5′-T of the dimer. Because a −1 frameshift was not observed in the replication of the DT_4 template, which could also form a similar slipped structure (i.e., 5′ for DT_4), this class of misaligned template-primers must not be readily elongated by the polymerase and either realigns itself or proceeds to a second slipped structure. The second slipped structure (4″) corresponds to a two-base-pair slip, where the 3′-end of the A_4 tract aligns with the T at the 5′-end of the T tract. Because this slipped intermediate has a normal Watson–Crick base-paired primer-terminus, further elongation (E4″) happens at a rate that is competitive with realignment (R4″), ultimately yielding a −2 frameshift product. To minimize interference with the polymerase-catalyzed primer-elongation step, it would make sense that the two extra Ts in the template would be situated as far from the primer-terminus as possible. Such a requirement is consistent with the observation that replication of the T_2DT_2, T_3DT, and T_4D templates does not result in any detectable −2 frameshift. In these cases, the extra Ts, or thymine dimer itself, would be too close to the primer-terminus

and might interfere with the primer-elongation step by reducing the binding constant of the polymerase for the template-primer. The DT_4 template is not expected to form a −2 frameshift despite the fact that it could accommodate the two displaced Ts, because the intermediate slipped structure (5″) would have a mismatched G·A primer-terminus. Purine–purine mispaired template-primer termini have been found to be poor substrates for further elongation by DNA polymerases (Joyce, 1989; Perrino & Loeb, 1989; Bebenek et al., 1990). Realignment of this mismatched and misaligned intermediate is therefore expected to take place in preference to further elongation.

Mechanism for the Formation of the −1 Frameshift Mutation. In competition with misalignment of the A_4 tract primer (step M4) in the bypass of the thymine dimer in TDT_3, the polymerase could elongate the primer by another A (step E4) leading to an aligned A_5 tract primer-template (5). Again, this intermediate could either be elongated or slip by one base pair. If slippage takes place, a primer-template (5′) is produced containing a normal Watson–Crick base-paired primer-terminus which can be elongated by the addition of C. Uninterrupted elongation of this intermediate (5C′) would lead to the observed −1 frameshift mutation. To obtain additional evidence for this process, intermediate 5 was prepared independently from the 15b-mer primer and the TDT_3 template (Figure 4). Elongation of this intermediate followed by *Eco*RI cleavage resulted in both a full-length product and a −1 frameshift product in amounts roughly equal to those observed in the elongation reaction with the 15a-mer primer (data not shown).

Mechanism for the Formation of the Base-Substitution Mutation. Competitive with the elongation of the primer terminating in C opposite the G in the TDT_3 template (5C′) would be realignment, leading to an aligned, but mismatched primer-template (5C). In this case, the cost of forming a mismatched primer-terminus is compensated by regaining a properly aligned template-primer. Further elongation of this template-primer would give the observed A → C transversion mutation. The mechanism whereby a substitution mutation results from transient misalignment is referred to as dislocation mutagenesis (Kunkel, 1985; Kunkel & Alexander, 1986) and is based on earlier proposals (Fresco & Alberts, 1960; Fowler et al., 1974) and experimental evidence (Kunkel, 1985). Additional evidence that this type of mechanism was operating in our system was obtained by independently preparing intermediate 5C by annealing the 16-mer primer to the TDT_3 template (Figure 4). Elongation of this intermediate led to both full-length and −1 frameshift products in approximately a 1:10 ratio (data not shown).

Dimer Site-Specificity of the Frameshift and Substitution Mutations. The failure to observe frameshift or substitution mutations in the replication of DT_4, T_2DT_2, T_3DT, and T_4D by exo⁻ T7 polymerase is consistent with the mechanism proposed for the origin of these mutations in the replication of TDT_3. None of the possible misaligned intermediates formed with this set of templates except one (vide infra) meets the structural requirements that appear to be necessary for elongation of a misaligned intermediate to be competitive with its realignment. These structural requirements are (1) a normal Watson–Crick base-paired primer-terminus and (2) the ability to position all extra nucleotides no closer than four nucleotides from the end of the primer-terminus. The only misaligned intermediate that meets these requirements is that in which an A_4 tract primer is misaligned with T_2DT_2 (4′, Figure 9). This intermediate is equivalent in structure to the

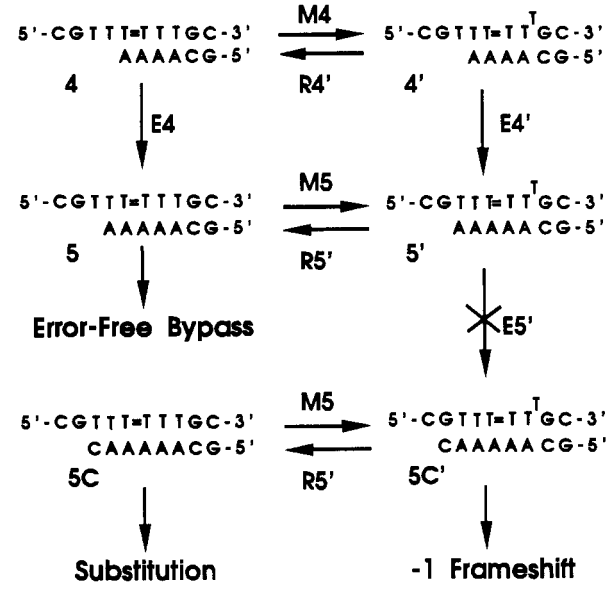


FIGURE 9: Proposed mechanistic scheme for the failure to observe mutations in the replicative bypass of the cis–syn thymine dimer of template T_2DT_2. See Figure 8 for an explanation of the symbols used.

intermediate leading to the −2 frameshift in TDT_3 (4″, Figure 8) except that there is only one extra nucleotide instead of two. A possible explanation for the failure of this intermediate to lead to a detectable amount of a −1 frameshift product in spite of what appears to be a normal primer-terminus and a suitable location for the displaced nucleotide rests on careful consideration of the other pathways competing with this one. One can argue that following an additional elongation step (E4′) realignment takes place (R5′) in preference to further elongation resulting in an aligned primer-template with a normal terminus (5). This is the same realignment step that competes for the −1 frameshift pathway in TDT_3 and results in the A → C mutation (R5C′, Figure 8) and would likewise diminish the amount of −1 frameshift that would occur in the replication of T_2DT_2. To gain additional support for this explanation, intermediate 5 (Figure 9) was independently prepared from the 15b-mer primer and the T_2DT_2 template (Figure 4) and subjected to elongation by Sequenase Version 2.0 followed by *Eco*RI cleavage. Under these conditions, only the correct-length bypass product was observed, and no product corresponding to a −1 frameshift was detected. If this is the case, one might ask why a −2 frameshift takes place at all in the replication of TDT_3. Unlike the corresponding intermediate in the replication of T_2DT_2, when intermediate 4C″ (Figure 10) realigns to give 4C′ or 4C, a mismatched C·T primer-terminus is formed. Hence, elongation of all of these realigned primer-templates is expected to be much slower than elongation of the properly matched, but misaligned, primer-template 4C″.

Inhibitory Effects of Misaligned DNA on Primer-Elongation. There are many proposals regarding the structure of misaligned DNA duplexes. NMR studies of synthetic models in solution suggest that an extra purine stacks with the flanking bases in a double helix (Hare et al., 1986; Woodson & Crothers, 1988). Comparative gel electrophoresis measurements suggest that an extra adenosine stacks within the helix and causes a bend of about 20°, primarily in the direction of tilt away from the extra base (Rice & Crothers, 1989). In contrast, crystallographic evidence indicates that the extra purine is looped out of the helix, leaving the structure of the double helix virtually unchanged (Joshua-Tor et al., 1988). In the case of extra pyrimidines, comparative gel electro-

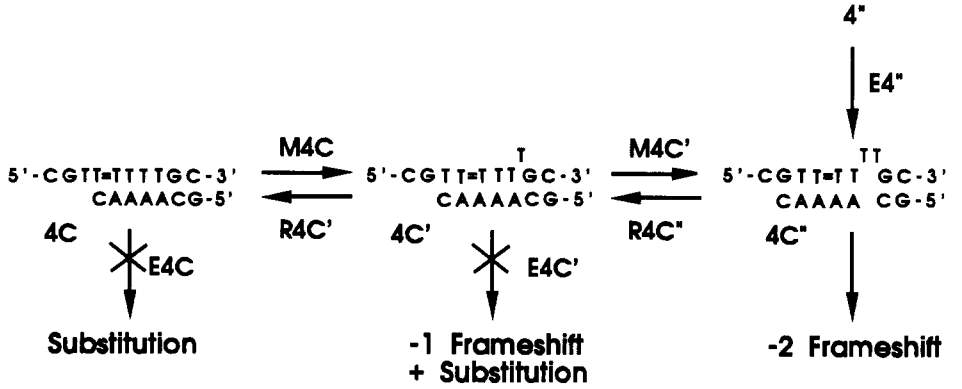


FIGURE 10: Proposed mechanistic scheme for the failure to observe mutations that might have been expected to occur along with the formation of the −2 frameshift mutation in the replicative bypass of TDT_3. See Figure 8 for an explanation of the symbols used.

phoresis measurements (Rice & Crothers, 1989) and NMR studies (Morden et al., 1983) indicate that the extra base spends most of the time outside the helix. Either class of structures could explain the inhibitory effect of an extra nucleotide on the polymerase-catalyzed elongation rate. The polymerase binding site for KF appears to be 5–9 base pairs long, and an extra-helical base within the binding site might be expected to have destabilizing steric interactions with the enzyme (Streisinger & Owen, 1985). Likewise, a bend induced by stacking of an extra nucleotide could also lead to similar destabilizing interactions with the enzyme.

In contrast to what is known about the structure of misaligned DNA in heterogeneous sequences, nothing is known about the structure of misaligned DNA involving damaged bases or involving homopolymer tracts. A thymine dimer places tremendous restraints on the structure of misaligned DNA if the extra nucleotide involved is one of the two that form the dimer and possibly one of the two that flank the dimer site. The fused structure of the dimer enforces a stacked-like structure to the two thymines of the dimer and might well restrict the orientation of the immediately flanking nucleotides. With regards to homopolymer tracts, A tracts adopt a conformationally distinct structure which is thought to be responsible for the bending of DNA caused by these structures (Crothers et al., 1990; Hagerman, 1990). Misaligned A tracts may have notably different structures from their heterogeneous counterparts, as the thymines in aligned A tracts have been found to form bifurcated hydrogen bonds with the complementary adenines (Nelson et al., 1987). Thus, A tracts appear to have an inherently slipped structure, and when misaligned they may be able to compensate for the effect of an extra nucleotide by way of their ability to engage in bifurcated hydrogen bonding. It is interesting to note that the conformational transition between the A tract structure and the B DNA structure appears to be around 37 °C (Chan et al., 1990), the temperature at which the amount of −2 frameshift begins to significantly diminish (Table II). With regard to the −2 frameshift mutations, little is known about the structure and properties of two extra bases in either heterogeneous or homopolymer tracts.

Role of the Polymerase in the Formation of Frameshift Mutations. To explain the high mutation frequency in the replication of TDT_3 by Sequenase Version 2.0 and its error-free replication by exo⁻ KF, one has to consider the rates of the various steps involved in the replication of the thymine dimer-containing template. As argued above, incorporation of an A opposite the T immediately to the 5′-side of the dimer (Figure 8, step E5) must proceed at a rate much slower than that for misalignment (step M5) to give intermediate 5′. This is the critical intermediate that ultimately leads to the −1 frameshift mutation provided that its realignment is slow relative to its elongation, a step that is expected to be polymerase-dependent.

The mechanisms by which misaligned template-primers are realigned during replication are possibly best understood for KF. The polymerase and 3′ → 5′ exonuclease activities of this polymerase reside on two distinct domains of 46 and 22 kDa, respectively (Freemont et al., 1986; Derbyshire et al., 1988). On the basis of the crystal structures of both duplex and single-stranded DNA bound to KF (Freemont et al., 1988), it was observed that a template-primer would have to slide a distance equivalent to about 8 base pairs to get from the polymerase site to the exonuclease site. At the same time, approximately 4 base pairs of the primer-terminus would be required to melt out to provide sufficient single-stranded DNA for binding to the exonuclease site. It has been proposed that this sliding and melting process might be an effective means for realigning intermediates that would otherwise lead to frameshift mutations (Kunkel, 1990). In support of this idea, the frequency of −1 frameshifts within runs of thymines was found to be higher for the single-domain 46-kDa polymerase of KF than for the two-domain 3′ → 5′ exo⁻ KF (Bebenek et al., 1990). Because the structure of the exonuclease-deficient polymerase was shown by crystallographic analysis to be virtually identical with the wild-type enzyme (Derbyshire et al., 1988), it seemed likely that sliding and melting of the template-primer could still take place and was responsible for lowering the frequency of frameshift mutations.

The sliding and melting mechanism would explain why replication of the dimer-containing T_6 tracts by exo⁻ KF did not lead to any detectable frameshift mutations. One can only speculate, however, on why replication of Sequenase Version 2.0 did lead to frameshift mutations. Sequenase Version 2.0 is a T7 DNA polymerase in which 28 amino acids in the 3′ → 5′ exonuclease activity domain have been deleted (Tabor & Richardson, 1989). Although little is known about the structure and mechanism of T7 polymerase and how they are affected by removing part of the exonuclease domain, one can imagine that sliding and melting of the template-primer does not take place with this exo-deficient polymerase as it does with exo⁻ KF. This could explain why elongation of a misaligned template-primer with a properly matched terminus such as intermediate 5′ of Figure 8 (step E5′) could compete effectively with realignment (step R5′). Further elongation would lead to further stabilization of the misaligned template-primer and ultimately to the observed −1 frameshift mutation following increasingly rapid elongation steps.

Another possible mechanism that could contribute significantly to the realignment of frameshift intermediates in replication by exo⁻ KF, but much less so by Sequenase Version

2.0, involves a reversible dissociation of the polymerase from the template-primer. This dissociative pathway was shown to compete with the intramolecular sliding pathway in the exonucleolytic editing of mismatched template-primers by KF (Joyce, 1989). T7 polymerase, on the other hand, is a highly processive enzyme and has a very high binding constant for template-primers (vide infra). In this case, realignment of a misaligned intermediate via a dissociation–reassociation pathway would be expected to be slow, thereby making the elongation step more competitive.

Relationship of This Study to Other Studies. There have been numerous studies of spontaneous [for a recent review, see Kunkel (1990)] and mutagen-induced [for recent examples, see Refolo et al. (1987), Sahm et al. (1989), Strauss (1989), and Schaaper et al. (1990)] frameshift mutations. Due to low mutation frequencies, all these studies have relied on in vivo systems to assay and quantify the replication errors. Because of this, the mechanisms responsible for the errors had to be inferred from the in vivo mutation data. Furthermore, in the mutagen-induced studies, heterogeneous substrates have invariably been used, making it virtually impossible to unambiguously assign a specific mutation to a specific adduct. Our system is unique in that pure, well-characterized, site-specific adducts are used as substrates and that the major mutagenic events are of such a high frequency that they can be directly quantified by a gel electrophoresis assay without resorting to biological amplification methods. These features make it possible for the first time to study the kinetics and mechanism of DNA adduct-induced frameshift mutagenesis in detail. We attribute the high frequency of frameshift mutations to three inherent properties of our system. First, the thymine dimer greatly attenuates the elongation steps catalyzed by Sequenase Version 2.0, thereby giving sufficient time for misalignment steps to occur. Second, the thymine dimer codes primarily as though it were an intact pair of thymines, ensuring that misaligned intermediates will be properly base-paired for efficient elongation. Third, the T tract is long enough so that the extra nucleotide in the misaligned intermediates can be located at a sufficient distance from the primer-terminus so as not to interfere with polymerase binding and subsequent elongation steps.

Implications for UV-Induced Mutagenesis in Vivo. Though the mechanism proposed for the origin of frameshift and substitution mutations described herein strictly only applies to Sequenase Version 2.0 in vitro, it may also be applicable to in vivo systems. In *E. coli*, UV-induced mutagenesis depends on the SOS response and is thought to involve pol III, RecA, umuC, and umuD′ proteins [for a review, see Echols and Goodman (1990)]. It has been proposed that RecA facilitates pol III-catalyzed elongation opposite a damage site by binding to the ϵ subunit of the polymerase and inhibiting its 3′ → 5′ exonuclease activity (Lu et al., 1986). Subsequent elongation past the damage site is then facilitated by the umuC and umuD′ proteins, possibly by increasing the binding constant of pol III for the damaged template (Bridges & Woodgate, 1985). Sequenase Version 2.0 mimics the properties of the SOS replication system by itself lacking a 3′ → 5′ exonuclease and by being highly processive. The high processivity facilitates elongation past the dimer and is due to the formation of a tight binding complex between the T7 gp5 protein and the *E. coli* thioredoxin protein. Thioredoxin increases the lifetime of T7 gp5 protein-bound template-primer by approximately 300-fold (Huber et al., 1987) and the processivity by approximately 1000-fold (Tabor et al., 1987). As such, this accessory protein functions in the manner postulated for the umuC and umuD′ proteins.

Evidence that replication of dimer-containing T tracts by Sequenase Version 2.0 is indeed a good model for what happens under SOS conditions comes from examining the most extensive data set for T tract deletions in *E. coli* (Miller, 1985). In that study, it was found that the longer the T tract, the higher the frequency of −1 frameshifts; specifically, 1 for a T_3 tract, 5 for a T_4 tract, and 11 for a T_5 tract. These data are consistent with our finding that only TDT_3 led to a significant amount of a −1 mutation, which we explained by the need to position the displaced nucleotide in a misaligned template-primer at some minimal distance from the primer-terminus. Intriguing as it is, this dependence of deletion frequency on T tract length has not been observed in other in vivo studies (LeClerc et al., 1984; Schaaper et al., 1987; Wood & Hutchinson, 1987). As further evidence, however, we have found three incidents of UV-induced substitution mutations at the 5′-end of T tracts that are likewise in accord with a dislocation mechanism, one in the *lacI* gene (Miller, 1985), one in the *lacZ* gene (LeClerc et al., 1984), and one in the λ cI gene (Wood & Hutchinson, 1987). In the first two cases the mutation was AAAAC → AAACC, and in the third it was AAAAAC → AAAACC. To our knowledge this is the first time that a dislocation mechanism has been invoked to explain the origin of a UV-induced mutation.

It is interesting to note that the cis–syn dimer that leads to the highest frequency of −1 frameshift mutations in vitro is not the one that is formed in highest frequency. We find that at 37 °C the ratio of DT_4:TDT_3:T_2DT_2:T_3DT:T_4D produced by irradiation of T_6 with 254-nm light is approximately 0.8:1.4:1.5:1.8:0.7 relative to cis–syn dimer formation at a nearby TT site in a heterogeneous sequence (data not shown). In general, dimers form in highest frequency in homopyrimidine tracts at the second site from the 3′-end (Brash & Haseltine, 1982). According to our mechanism, dimers formed at the 3′-end would be less likely to result in −1 frameshift mutations than dimers formed at the 5′-end, supporting the idea that photoproduct formation hotspots are not necessarily mutation hotspots (Brash et al., 1987). Despite the high yield of cis–syn dimer formation within T tracts under direct irradiation at 254 nm, there is some evidence to indicate that they may not be the principal cause of frameshift mutations. In a limited data set it was found that −1 frameshifts at T tracts were formed with higher frequency upon irradiation with 254 nm, than upon acetophenone-sensitized photolysis (Wood et al., 1984), conditions that almost exclusively produce cis–syn thymine dimers. To determine whether cis–syn thymine dimer-containing T tracts do indeed lead to frameshift and substitution mutations in vivo, we are currently incorporating the templates used in this study site-specifically into the *lacZ* gene of an M13mp18-derived bacteriophage. Repair-deficient *E. coli* hosts will then be transfected with these site-specifically photodamaged bacteriophages, and their progeny will be screened for mutants and sequenced.

Acknowledgments

We thank S. Lloyd (Vanderbilt University) for a generous gift of T4 *denV* endonuclease V and A. Sancar (University of North Carolina) for a generous gift of *E. coli* photolyase.

References

Ananthaswamy, H. N., & Pierceall, W. E. (1990) *Photochem. Photobiol. 52*, 1119–1136.

Banerjee, S. K., Christensen, R. B., Lawrence, C. W., & LeClerc, J. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8141–8145.

Bebenek, K., & Kunkel, T. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4946–4950.
Bebenek, K., Joyce, C. M., Fitzgerald, M. P., & Kunkel, T. A. (1990) *J. Biol. Chem. 265*, 13878–13887.
Brash, D. E., & Haseltine, W. A. (1982) *Nature 298*, 189–192.
Brash, D. E., Seetharam, S., Kraemer, K. H., Seidman, M. M., & Bredberg, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3782–3786.
Bridges, B. A., & Woodgate, R. (1985) *Mutat. Res. 150*, 133–139.
Cadet, J., & Vigny, P. (1990) in *Bioorganic Photochemistry*, (Morrison, H., Ed.) Vol. 1, Chapter 1, pp 1–272, John Wiley & Sons, New York.
Chan, S. S., Breslauer, K. J., Hogan, M. E., Kessler, D. J., Austin, R. H., Ojemann, J., Passner, J. M., & Wiles, N. C. (1990) *Biochemistry 29*, 6161–6171.
Clark, J. M., Joyce, C. M., & Beardsley, G. P. (1987) *J. Mol. Biol. 198*, 123–127.
Crothers, D. M., Haran, T. E., & Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093–7096.
Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M., & Steitz, T. A. (1988) *Science 240*, 199–201.
Echols, H., & Goodman, M. F. (1990) *Mutat. Res. 236*, 301–311.
Fowler, R. G., Degnen, G. E., & Cox, E. C. (1974) *Mol. Gen. Genet. 33*, 179–191.
Freemont, P. S., Ollis, D. L., Steitz, T. A., & Joyce, C. M. (1986) *Proteins: Struct., Funct., Genet. 1*, 66–73.
Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., & Steitz, T. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8924–8928.
Fresco, J. R., & Alberts, B. M. (1960) *Proc. Natl. Acad. Sci. U.S.A. 46*, 311–321.
Hagerman, P. J. (1990) *Annu. Rev. Biochem. 59*, 755–781.
Hare, D., Shapiro, L., & Patel, D. J. (1986) *Biochemistry 25*, 7456–7464.
Hsia, H. C., Lebkowski, J. S., Leong, P.-M., Calos, M. P., & Miller, J. H. (1989) *J. Mol. Biol. 205*, 103–113.
Huber, H. E., Tabor, S., & Richardson, C. C. (1987) *J. Biol. Chem. 262*, 16224–16232.
Hutchinson, F. (1987) *Photochem. Photobiol. 45*, 897–903.
Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appella, E., & Sussman, J. L. (1988) *Nature 334*, 82–84.
Joyce, C. M. (1989) *J. Biol. Chem. 264*, 10858–10866.
Kemmink, J., Boelens, R., Koning, T., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1987) *Nucleic Acids Res. 15*, 4645–4653.
Kunkel, T. A. (1985) *J. Biol. Chem. 260*, 5787–5796.
Kunkel, T. A. (1990) *Biochemistry 29*, 8003–8011.
Kunkel, T. A., & Alexander, P. S. (1986) *J. Biol. Chem. 261*, 160–166.
LeClerc, J. E., Istock, N. L., Saran, B. R., & Allen, R. J. (1984) *J. Mol. Biol. 180*, 217–237.
Lu, C., Scheuermann, R. H., & Echols, H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 619–623.
Miller, J. H. (1985) *J. Mol. Biol. 182*, 45–68.
Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I., Jr. (1983) *Biochemistry 22*, 5557–5563.
Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221–226.
Nguyen, H. T., & Minton, K. W. (1988) *J. Mol. Biol. 200*, 681–693.
Perrino, F. W., & Loeb, L. A. (1989) *J. Biol. Chem. 264*, 2898–2905.
Refolo, L. M., Bennett, C. B., & Humayun, M. Z. (1987) *J. Mol. Biol. 193*, 609–636.
Rice, J. A., & Crothers, D. M. (1989) *Biochemistry 28*, 4512–4516.
Ripley, L. S. (1990) *Annu. Rev. Genet. 24*, 189–213.
Sahm, J., Turkington, E., LaPointe, D., & Strauss, B. (1989) *Biochemistry 28*, 2836–2843.
Schaaper, R. M., Dunn, R. L., & Glickman, B. W. (1987) *J. Mol. Biol. 198*, 187–202.
Schaaper, R. M., Koffel-Schwartz, N., & Fuchs, R. P. P. (1990) *Carcinogenesis 11*, 1087–1095.
Strauss, B. S. (1989) *Ann. Ist. Super. Sanita 25*, 177–189.
Streisinger, G., & Owen, J. E. (1985) *Genetics 109*, 633–659.
Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., & Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77–84.
Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767–4771.
Tabor, S., & Richardson, C. C. (1989) *J. Biol. Chem. 264*, 6447–6458.
Tabor, S., Huber, H. E., & Richardson, C. C. (1987) *J. Biol. Chem. 262*, 16212–16223.
Taylor, J.-S. (1990) *J. Chem. Educ. 67*, 835–841.
Taylor, J.-S., & O'Day, C. L. (1990) *Biochemistry 29*, 1624–1632.
Taylor, J.-S., Brockie, I. R., & O'Day, C. L. (1987) *J. Am. Chem. Soc. 109*, 6735–6742.
Taylor, J.-S., Garrett, D. S., Brockie, I. R., Svoboda, D. L., & Telser, J. (1990) *Biochemistry 29*, 8858–8866.
Wang, C.-I., & Taylor, J.-S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9072–9076.
Wood, R. D., & Hutchinson, F. (1984) *J. Mol. Biol. 173*, 293–305.
Wood, R. D., & Hutchinson, F. (1987) *J. Mol. Biol. 193*, 637–641.
Wood, R. D., Skopek, T. R., & Hutchinson, F. (1984) *J. Mol. Biol. 173*, 273–291.
Woodson, S. A., & Crothers, D. M. (1988) *Biochemistry 27*, 3130–3141.