688–696 Nucleic Acids Research, 2008, Vol. 36, No. 2 Published online 13 December 2007
doi:10.1093/nar/gkm1089

DNA polymerase kappa produces interrupted mutations and displays polar pausing within mononucleotide microsatellite sequences

Suzanne E. Hile and Kristin A. Eckert*

Department of Pathology, Gittlen Cancer Research Foundation, The Pennsylvania State University College of Medicine, 500 University Drive, Hershey, PA 17033, USA

Received October 18, 2007; Revised and Accepted November 19, 2007

ABSTRACT

Microsatellites are ubiquitously present in eukaryotic genomes and are implicated as positive factors in evolution. At the nucleotide level, microsatellites undergo slippage events that alter allele length and base changes that interrupt the repetitive tract. We examined DNA polymerase errors within a $[T]_{11}$ microsatellite using an *in vitro* assay that preferentially detects mutations other than unit changes. We observed that human DNA polymerase kappa (Pol κ) inserts dGMP and dCMP within the $[T]_{11}$ mononucleotide repeat, producing an interrupted 12-bp allele. Polymerase β produced such interruptions at a lower frequency. These data demonstrate that DNA polymerases are capable of directly producing base interruptions within microsatellites. At the molecular level, expanded microsatellites have been implicated in DNA replication fork stalling. Using an *in vitro* primer extension assay, we observed sequence-specific synthesis termination by DNA polymerases within mononucleotides. Quantitatively, intense, polar pausing was observed for both pol κ and polymerase α-primase within a $[T]_{11}$ allele. A mechanism is proposed in which pausing results from DNA bending within the duplex stem of the nascent DNA. Our data support the concept of a microsatellite life-cycle, and are consistent with the models in which DNA sequence or secondary structures contributes to non-uniform rates of replication fork progression.

INTRODUCTION

Approximately 3% of the human genome contains microsatellite DNA sequences, which are present on every chromosome at an average density of ~14 000 bp/Mbp (1). Therefore, repetitive microsatellite DNA comprises a significant component of genome replication that must occur faithfully each cell cycle. Direct sequence analyses of eukaryotic genomes have revealed that the precise composition of microsatellites is heterogeneous, ranging from pure arrays of a single repetitive sequence, to complex arrays containing several types of repetitive units, to arrays that are interrupted by single base changes or insertion/deletion mutations (2). Sequence variation in common microsatellites has been proposed to have a positive role in evolution (3), and numerous reports have illustrated the phenotypic (gene expression) effects of microsatellite length variation (4,5). Evolutionary models of microsatellites propose that length distributions reflect a balance between expansion or contraction errors and point mutations within the allele (6). The accumulation of base interruptions has been proposed to break the repetitive array, resulting in death of the microsatellite (7). Expansion/contraction mutations are generally assumed to occur by slipped strand mispairing (8), and both *in vivo* and *in vitro* data are consistent with microsatellite errors that result in the gain or loss of repeat units (9–11).

Microsatellites also have been implicated in genome stability at the molecular level, through functional effects on chromatin organization, recombination and DNA replication (12). Expanded trinucleotide alleles can adopt non-B DNA secondary structures, thereby causing DNA polymerase pausing *in vitro* and replication fork pausing *in vivo* (13,14). Replication fork arrest may precede replication fork collapse, resulting in double strand breaks and genome rearrangements (15,16). In contrast to the trinucleotide microsatellites, little is known about the biochemistry of DNA replication through the more highly abundant, shorter, mono- and dinucleotide repetitive sequences. Like the trinucleotide repeats, these sequences also have the potential to adopt non-B form DNA structures (17). We previously observed strong DNA polymerase α-primase (pol α-primase) pauses within a $[TC]_{20}$ dinucleotide allele that were caused by triplex DNA formation between the nascent duplex DNA and the ssDNA template during DNA synthesis (18).

*To whom correspondence should be addressed. Tel: +1 717 531 4065; Fax: +1 717 531 5634; Email: kae4@psu.edu

The identification of DNA polymerases present at the human replication fork is currently under intense investigation (19,20). Current models to explain resolution of stalled replication forks invoke replacement of replicative polymerases or gap-filling behind the replication fork by specialized DNA polymerases (21–23). DNA polymerase κ (Pol κ) has been detected in nuclear foci at a low level in undamaged human cells (24), suggesting pol κ may be active at replication forks. Pol κ also is capable of performing efficient extension synthesis from mispaired and misaligned template-primer termini *in vitro* (25,26). Because human polymerases differ significantly in fidelity (27,28), the identity of the polymerase(s) synthesizing nascent microsatellite DNA is expected to impact microsatellite stability. In the current study, we investigated the biochemistry and accuracy of pol κ during DNA synthesis through mononucleotide alleles. Our mutagenesis studies show that DNA polymerases may influence the evolution of microsatellites by creating base interruptions, thereby rendering the microsatellite sequences more stable. Our DNA synthesis progression studies show that short microsatellites also may inhibit DNA polymerases by forming alternative DNA secondary structures.

MATERIALS AND METHODS

Reagents and vectors

Purified full-length (99 kDa) human pol κ was kindly supplied by Dr Zhigang Wang (University of Kentucky, Lexington, KY) or was purchased from Enzymax (Lexington, KY). Purified human pol α-primase complex was purchased from Chimerx (Madison, WI). The 3′→5′ exonuclease-deficient form of *Escherichia coli* DNA polymerase I large fragment (D424A) was a generous gift from Dr Catherine Joyce (Yale University). Recombinant DNA polymerase β was purified as described (29). The mononucleotide vectors containing $[T/A]_9$ (inserted between bases 111 and 112 of the target HSV-tk gene) or $[G/C]_9$ (inserted between bases 110 and 111) inserts were constructed by cloning products of an oligonucleotide-directed *in vitro* T7 DNA polymerase reaction as described (30). Bases flanking the inserts were included as part of the microsatellite region in data analyses when appropriate. Therefore, the total allele length of the microsatellite is always reported. The microsatellite unit located closest to the 5′ end of the HSV-tk sense strand was denoted as unit 1.

In vitro HSV-*tk* mutagenesis assay

Linear DNA fragments and ssDNA were prepared and used to construct gapped duplex (GD) molecules, as described (31), with the following modification. A StuI restriction site was created at HSV-*tk* position 180, and subcloned into the pSS vector (11). Restriction enzyme digestion with MluI and StuI and subsequent hybridization to ssDNA creates a hetereoduplex molecule containing an 81 nt gap that encompasses the microsatellite allele, but excludes the majority of the HSV-*tk* coding sequence. Oligonucleotide-primed ssDNA templates were constructed to initiate DNA synthesis at position 169 of the HSV-*tk* gene. The *in vitro* reactions contained 1 pmol of template DNA at 40 nM concentration. Reaction conditions for pol β were as described (11), except that 10 pmol enzyme were used. Reaction conditions for pol κ were as below, except that 8 pmol enzyme were used. To sample reaction products for mutations, small fragments were prepared by MluI and StuI restriction digestion and hybridized to the corresponding GD as described (11,31). Successful hybridization to GD was achieved for all reactions, as determined by agarose gel electrophoresis. To select for HSV-*tk* mutations, an aliquot of DNA from the final hybridization was used to transform *rec*A13, *upp, tdk E. coli* strain FT334 by electroporation, and plated on VBA selective media (31). Bacteria harboring HSV-*tk* mutant plasmids were selected by plating in the presence of 40 μM 5-fluoro-2′-deoxyuridine and 50 μg/ml chloramphenicol. The DNA sequence of the HSV-*tk* gene in the MluI-StuI region of independent mutants was determined as described (11). Differences in proportions of specific types of mutations were analyzed statistically using Fisher's exact test (two-tailed).

Polymerase pausing

Pausing analyses were carried out as described (18). Briefly, DNA synthesis templates were created by hybridizing a [γ-^{32}P] 5′-end labeled oligonucleotide to the appropriate ssDNA at a 1:1 molar ratio. The standard reactions for pol κ primer extension contained 25 mM potassium phosphate at the indicated pH, 5 mM $MgCl_2$, 5 mM dithiothreitol, 100 μg/ml BSA, 1 mM dNTPs and 200 fmol primer-template. Synthesis was initiated upon addition of 50–100 fmol pol κ. Under these conditions, we routinely observe reaction products up to 100 nt or more in length, and the distribution of these products is independent of reaction time. Conditions for the pol α-primase reactions were as described (18). Analyses to test triplex potential were done analogously, substituting the nucleotide analog, 7-deaza-dATP, for dATP. Quantitation of reaction products was performed using a Molecular Dynamics Phosphorimager and ImageQuant software (Sunnyvale, CA). Total percent extension is the amount of total extended molecules (corrected for percent hybridization and background) divided by this number plus the amount of corrected primer molecules. The percent synthesis termination at a particular nucleotide was calculated by dividing the number of reaction product molecules of that length by the total number of template-primer molecules extended in the reaction.

RESULTS

Polymerase mutagenesis within a $[T]_{11}$ mononucleotide allele

The sequence composition of microsatellites is non-random, and in the human genome, poly(dA/dT) mononucleotides are the most abundant class (1,32). In this study, we set out to determine mutational mechanisms within mononucleotide alleles using purified human DNA polymerases. We previously developed an *in vitro* assay to measure DNA polymerase errors within microsatellite alleles, using the HSV-*tk* gene as a mutational

Table 1. Mutational specificity of DNA polymerases κ and β

Region/mutational class[a]	Number of events (Proportion of region)	
	pol κ	pol β
T_{11} microsatellite	28	16
T deletion[b]	0	12 (0.75)
T insertion[c]	4 (0.14)	1 (0.06)
G insertion	20 (0.71)	2 (0.12)
C insertion	2 (0.07)	0
Base substitution[d]	2 (0.07)	1 (0.06)
HSV-tk coding	65	24
Base substitutions	9 (0.14)	3 (0.12)
Frameshifts[e]	53 (0.82)	19 (0.79)
Complex	3 (0.05)	2 (0.08)
HSV-tk$^-$ Mutant Frequency[f]	24×10^{-4}	35×10^{-4}

[a]Mutations that create alleles longer than T_9 do not result in a mutant phenotype due to transcriptional slippage within the microsatellite.
[b]Includes −2T errors, −1T occurring as a multiple mutation, and complex errors involving the microsatellite.
[c]Includes + 1T errors occurring as a multiple mutation and complex errors involving the microsatellite.
[d]Base substitution errors within the microsatellite occurred as multiple mutations.
[e]All frameshifts were deletion errors for both polymerases.
[f]The ssDNA mutant frequency is 1.3×10^{-4} for this template.

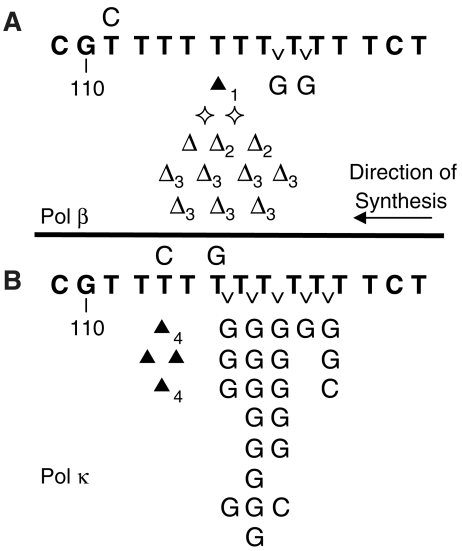

Figure 1. Mutational specificity of DNA polymerases in the $[T]_{11}$ assay. Middle sequence is the microsatellite target. Base substitutions are indicated above the sequence and frameshifts are indicated below. Symbols: open triangle: one T deletion (paired with a large 28bp coding region deletion); filled triangle: one T insertion (paired with a C to T at 100 or a C to T at 130); open diamond: two T deletion; letter 'v': insertion of noncanonical base. Subscripts indicate multiple errors detected within a single mutant. 1, T insertion paired with G to T at 110; 2, T deletion paired with T to C at STR; 3, T deletion paired with C 109 deletion; 4, T insertion paired with G insertion.

reporter (11). In the published system, polymerase errors are produced within either an artificial in-frame microsatellite sequence or the HSV-*tk* coding region, and are detected phenotypically using selective plating of *E. coli*. In our current study, we modified our existing assay by excluding ~100 nt of target template sequence within the coding region in order to improve detection of polymerase errors within the microsatellite allele. Our first modified HSV-*tk* construct contained an in-frame $[T]_{11}$ microsatellite (constructed by inserting a T_9 sequence within an existing TT sequence of the HSV-tk gene). The major types of cellular mutational events at mononucleotide alleles are insertions or deletions of one repeat unit (33–35). However, in *E. coli*, out-of-frame mutations that occur in repetitive poly(dA) or poly(dT) sequences longer than 9 nt can be placed back in-frame by transcriptional slippage during RNA polymerase elongation (36). In agreement with this phenomenon, we have confirmed that out-of-frame $[T]_{10}$ and $[T]_{19}$ alleles do not produce an HSV-*tk* mutant phenotype, whereas an out-of-frame $[T]_7$ mutant was detectable (data not shown). Thus, our modified mutational assay is biased against detection of one unit changes within our $[T]_{11}$ allele, allowing improved detection of non-canonical mutations.

We used this specialized mutagenesis assay to determine the frequency and specificity of errors produced by human DNA polymerases κ and ß. For pol κ, we observed that the majority of HSV-*tk* mutants contained errors within the coding region (Table 1). Thirty percent of pol κ errors were found within the $[T]_{11}$ microsatellite sequence. Of these, the majority (78%) of errors observed are insertion of a dGMP or dCMP residue within the $[T]_{11}$ microsatellite sequence (Table 1), corresponding to an error frequency of 5.6×10^{-4}. The pol κ insertions are not evenly distributed throughout the microsatellite allele, but primarily occur midway during DNA synthesis, between positions T6 and T7 or between T7 and T8 (Figure 1). These alleles produce new out-of-frame microsatellite loci (for example, T_6GT_5) that are detectable mutants. DNA synthesis by pol β produced a higher proportion of errors (40%) within the $[T]_{11}$ microsatellite sequence, relative to pol κ. In contrast to pol κ, however, the majority of pol β mutants are either deletion of one T nucleotide within the microsatellite allele (detected as a multiple mutation) or complex errors involving a T deletion within the microsatellite (see legend to Figure 1). Only 12% of pol ß microsatellite errors are non-iterative base insertions, corresponding to a frequency of 1.8×10^{-4} (Table 1). Thus, pol κ and pol β show a statistically significant difference in the proportion of insertion errors within the microsatellite ($P < 0.0001$, Fisher's exact test). The mutational specificity difference between the polymerases is specific to the microsatellite region, as the two polymerases produce a similar proportion and type of frameshift errors within the coding region ($P = 0.77$, Fisher's exact test) (Table 1).

DNA synthesis termination within a $[T]_{11}$ mononucleotide allele

In addition to nucleotide sequence changes that result in allelic polymorphism, microsatellite sequences can adopt non-B DNA forms that inhibit DNA synthesis. To examine whether DNA synthesis is inhibited within short mononucleotide alleles, we quantitated pol κ

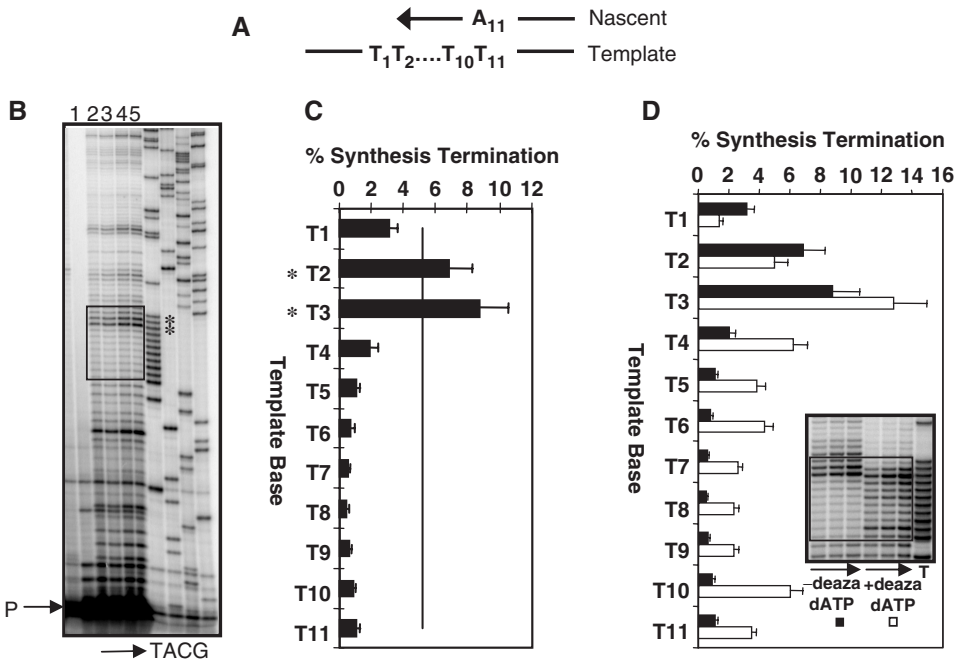

Figure 2. Polymerase κ termination within the T_{11} microsatellite. (**A**). Cartoon depicting DNA synthesis through the allele. (**B**). Representative phosphorimager scan of primer extension reaction products. Reactions were incubated at pH 7.0 and 30°C. Arrow indicates reaction time from 5 to 60 min. P, unextended primer. Lane 1, no polymerase control. Lanes 2–5, products of 5, 15, 30 and 60 min reactions. The 30 min reaction gave a total percent extension of 16.5%. Boxed region is the microsatellite. (**C**) Quantitation (30 min) for each template position in the microsatellite (mean of seven reactions). The continuous line represents the value of the percent synthesis termination that is 1 SD above the mean of all values within a given microsatellite. Asterisks indicate pause sites. (**D**) Deaza-dATP effect on termination. Quantitation of 30 min reactions without deaza-dATP (filled bars) or with deaza-dATP (open bars, three independent reactions). Inset: representative scan of primer extension products, with arrows indicating reaction time (5–30 min) and a box to indicate the beginning and end of the microsatellite sequence.

pausing within the $[T]_{11}$ sequence using an *in vitro* primer extension assay under standard reaction conditions of 30°C and pH 7.0 (Figure 2). Termination profiles were determined empirically using enzyme to template ratios that limit polymerase rebinding to previously extended DNA products. For example, we observed 16% total primer-template extension after 30 min (lane 4) and this distribution is independent of time (lanes 2–5). Similar results were obtained for total percent extension values between 5 and 40%; therefore, comparisons of subsequent quantitative analyses within the 5′-[STR]-3′ templates were made using reactions with a percent extension within this range. For microsatellites within the complementary strand, we note that a higher total percent extension by pol κ is necessary to achieve enough synthesis through the microsatellite for quantitative analyses, as shown previously for pol α-primase (18).

The percent synthesis termination was quantitated for each nucleotide within the $[T]_{11}$ microsatellite (Figure 2C). The overall mean percent termination for the 11 nt within this microsatellite was 2.40 ± 0.43 (Table 2). A template position was considered a polymerase pause site if the percent synthesis termination for that residue was 1 SD above the mean of all microsatellite sites (shown as a continuous line in Figure 2C). We observed two prominent pol κ pauses at the 3′ end of the nascent strand, at residues T2 and T3. The sites of intense pausing (Figure 2) are not related to the sites of non-iterative base insertions (Figure 1). To examine whether the pronounced pausing within poly(dT) microsatellite sequences is polymerase-specific, we also examined pausing by the replicative polα-primase complex under its standard reaction conditions of 37°C and pH 7.5. A polar pausing pattern again was observed for pol α-primase synthesis. Pol α-primase

Table 2. Polymerase κ pausing profile at mononucleotide microsatellite sequences

STR sequence context	pH	Temperature (°C)	Mean percent termination per nt[a]
5′-[STR]-3′			
T_{11}	6.5	30	2.39 (2)
	7.0	30	2.40 ± 0.43 (7)
	7.0 + deaza-dA[b]	30	4.57 ± 0.66 (3)
	7.0	35	1.08 (2)
	7.0	40	0.49 (2)
G_{10}	7.0	30	1.29 ± 0.07 (4)
	7.0	35	1.22 (2)
	7.0	40	1.00 ± 0.13 (3)
A_{11}	7.0	30	1.28 ± 0.06 (5)
3′-[STR]-5′			
C_{10C}	7.0	30	1.31 ± 0.14 (3)
A_{11C}	7.0	30	0.51 ± 0.15 (5)

[a]Data are the mean or the mean ± SD for the number of trials given in parentheses. Percent termination per nt is the amount of synthesis termination within the entire microsatellite, relative to total synthesis, normalized for the number of nucleotides in the allele.
[b]Reactions contained 7-deaza-dATP in place of dATP.

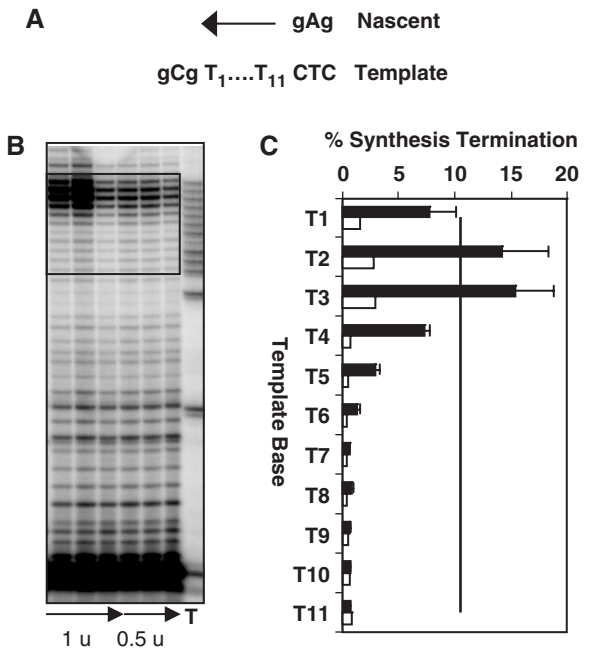

Figure 3. Polymerase α-primase termination within the T_{11} microsatellite. (**A**) Cartoon depicting DNA synthesis through the allele. (**B**) Representative phosphorimager scan of primer extension reaction products. Pol α reactions contained either 0.5 or 1 unit of enzyme and were incubated at 37°C. Arrow indicates increase in reaction time from 5 to 30 min. (**C**) Quantitation of the percent synthesis termination (30 min) for each template position in the microsatellite for pol α-primase at 37°C (filled bars, three experiments) and pol κ at 35°C (open bars, two experiments).

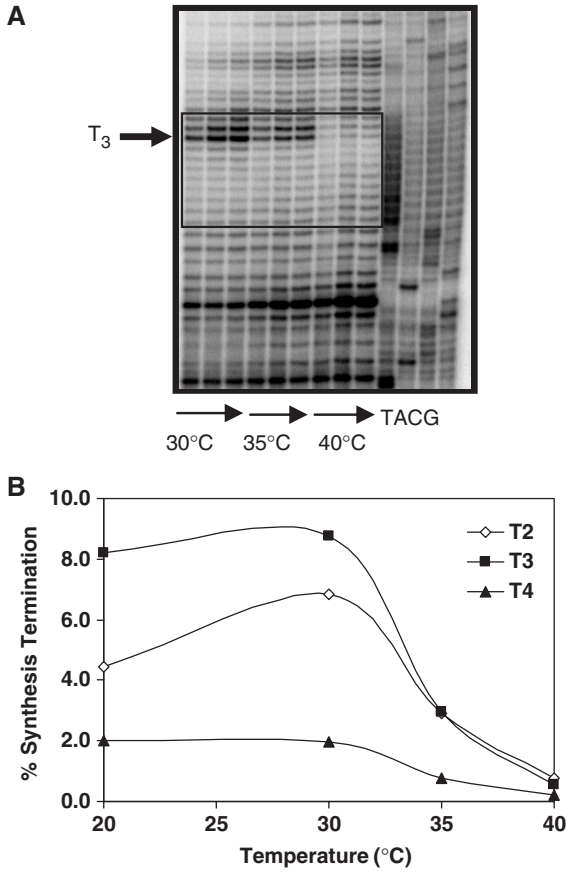

Figure 4. Effect of reaction temperature on polymerase κ termination at the T_{11} allele. (**A**) Representative phosphorimager scan of primer extension reaction products. Arrows indicate reaction time (5–30 min). Boxed area shows the beginning and end of the microsatellite sequence. (**B**) Variation of percent synthesis termination with temperature for nucleotides T2, T3 and T4. Each point represents mean of two independent experiments (30 min).

displayed a mean percent termination per nucleotide of 4.78 ± 0.88, ~4-fold greater than pol κ at a similar reaction temperature of 35°C (Table 2). Although reaction conditions were not identical, termination by polα-primase within residues T1 through T5 was 5–9-fold more intense than was pol κ (Figure 3C).

Mechanism of polar pausing within the $[T]_{11}$ microsatellite

Poly(dA/dT) sequences have the potential to form intermolecular triplexes *in vitro* (37). We directly tested whether intramolecular triplex DNA structures were involved in polar pausing. The high pol κ termination probability within the $[T]_{11}$ allele was not affected by lowering the reaction pH to 6.5 (Table 2), or by replacing dATP with 7-deaza-dATP (Figure 2D, Table 2). The enhanced termination observed upon addition of 7-deaza-dATP is consistent with published reports showing a decreased catalytic efficiency for incorporation of 7-deaza-dATP, relative to dATP, by pol κ (38). We observed strong, polar pauses at T1 and T2 by exonuclease-deficient Klenow polymerase; the pauses were somewhat reduced, but not eliminated, by addition of deaza-dATP (Supplementary data, Figure S1). We next examined whether the pol κ pause sites are temperature-dependent. We observed a 5-fold decrease in overall pol κ termination within the $[T]_{11}$ allele as the reaction temperature was increased from 30°C to 40°C (Table 2), and the strong pause sites at positions T2 and T3 were eliminated at the higher reaction temperature (Figure 4A). Site-specific quantitation revealed that the percent synthesis termination decreased 9–15-fold as the temperature increased from 30°C to 40°C (Figure 4B).

To determine whether the pronounced pausing within the $[T]_{11}$ allele is sequence-dependent, we quantitated the pausing profiles along an $[A]_{11}$ and a $[G]_{10}$ mononucleotide allele within the same sequence context, and $[A]_{11C}$ and $[C]_{10C}$ alleles within the complementary sequence context (Figure 5). Data are graphed using the same scale as that for the T_{11} allele, in order to display the stark difference between pol κ termination at the $[T]_{11}$ sequence versus the other mononucleotides. The mean percent termination per nucleotide within the $[A]_{11}$ allele was 1.28 ± 0.06, while that within the $[G]_{10}$ allele was 1.29 ± 0.07, both values approximately 2-fold lower than the $[T]_{11}$ allele (Table 2). We also calculated the termination probability (39) within each allele in order to control for potential differences in the extent of polymerase termination within sequences prior to the microsatellites. Comparison of the termination probability for pol κ synthesis within the $[T]_{11}$ versus the $[A]_{11}$ allele

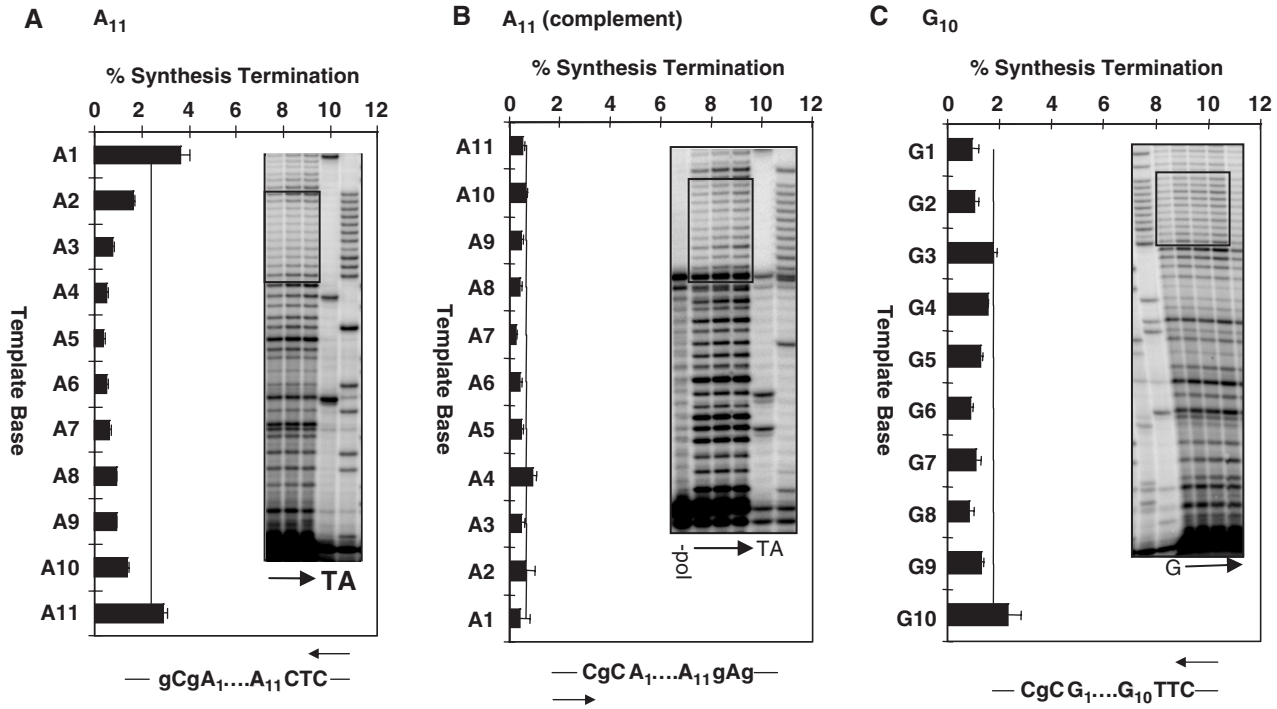

Figure 5. Sequence dependence of pol κ termination within mononucleotide alleles. Phosphorimager quantitation of percent synthesis termination for 30 min reactions at pH 7.0 and 30°C are shown for (**A**). the A_{11} allele (mean of five experiments); (**B**) the A_{11} allele in the complementary sequence context (mean of five experiments); and (**C**) the G_{10} allele (mean of four experiments). Insets show representative phosphoimager scans of primer extension products, with arrows indicating reaction time (5–30 min). Boxed areas indicate the beginning and end of each microsatellite allele. Sequence below graph indicates sequence context of microsatellite and arrows indicate direction of synthesis. –pol, no polymerase control, showing that the dark band at the 3′ end of the A_{11C} allele is an impurity in the primer and not a pause site.

again demonstrates that pol κ displays enhanced pausing within the poly(dT) sequence (Supplementary data, Figure S2). In addition, we examined the $[A]_{11C}$ allele in which the repetitive A tract is in the opposite sequence context, complementary to that of the $[T]_{11}$ allele. In this case, termination at the $[A]_{11C}$ allele was very weak, 5-fold lower than that of $[T]_{11}$ (Figure 5B, Table 2). In contrast to the complementary $[T/A]_{11}$ pair, the mean percent termination per nucleotide was the same along the complementary $[G]_{10}$ and $[C]_{10C}$ alleles (Table 2).

DISCUSSION

Production of mononucleotide interruptions by DNA polymerases

We used a specialized microsatellite mutagenesis assay to determine the frequency and specificity of DNA pol κ and pol β errors within a $[T]_{11}$ microsatellite. Strikingly, we observed that pol κ microsatellite errors involved the insertion of a dGMP or dCMP residue approximately midway within the $[T]_{11}$ microsatellite sequence (Figure 1). These insertion events were specific for pol κ, as they were observed infrequently with pol β, and for the microsatellite, as pol κ did not produce such insertions within the downstream HSV-*tk* coding region (Table 1). These results demonstrate that DNA polymerases *in vitro* can directly create interrupted alleles from a pure microsatellite. Interrupted microsatellite alleles have been identified in phylogenetic genomic comparisons and may be responsible for microsatellite allele shortening (40). For example, a single base interruption of a poly(dT) locus was observed among primates lineages that produced a monomorphic, interrupted allele from a polymorphic, pure allele (41). In yeast, base interruptions of microsatellites result in increased genetic stability, relative to pure alleles (42). We suggest DNA synthesis by pol κ will promote mononucleotide stability because a major type of error it produces (interruptions) within the microsatellite are protective mechanistically.

DNA polymerase production of interrupted alleles occurred primarily through insertion of a non-iterative base within the repetitive microsatellite sequence, rather than by a base substitution mutation within the allele (Figure 1). Insertion of a non-iterated base within a repetitive sequence is extremely unusual in the literature of DNA polymerase fidelity, but has been occasionally detected by pol ζ (43), pol λ (44), pol η (45) and by pol ß using a damaged DNA template (46). Polymerase one-base insertion errors typically involve the insertion of the same nucleotide within a repetitive DNA sequence, or insertions of a base within a non-repetitive sequence (27). Several mechanisms for the *in vitro* production of misalignment-mediated DNA synthesis errors by purified DNA polymerases have been described (27,28). Deletions at non-iterative sequences have been proposed to be

initiated by mispairing, followed by a template-primer misalignment. We propose that the pol κ dGMP insertions within the $[T]_{11}$ microsatellite are initiated by formation of a T–dCMP mispair. Previous fidelity studies of pol κ have shown that the most frequently formed base substitution by pol κ is a T → G error, consistent with a high rate of T–dCMP mispairing (47,48). The highly repetitive nature of the $[T]_{11}$ template sequence may facilitate rearrangement of the nascent DNA containing the dCMP to a misaligned intermediate by slippage elsewhere in the repetitive tract. In support of this model, we observe single T→G base substitutions within the microsatellite allele, which is the expected outcome in the absence of the misalignment step. An alternative model to explain the production of insertion mutations is transient misalignment (27), in which the primer strand relocates from position T6-T10 (the point of insertion) to position T1, with a concomitant template loop out of 4–8 bases. In this model, the dCMP insertion is templated by the G at position 110, followed by realignment of the primer strand. These two models can be differentiated experimentally by examining the effects of 5′ sequence context on the specificity of interruptions. Importantly, either model proposes that strand slippage within the repetitive microsatellite is involved in the production of interrupted alleles.

DNA structure and polymerase pausing

The study of replication dynamics through common microsatellite sequences and the key polymerases involved is crucial for understanding genome stability. We previously reported that the extent of DNA synthesis termination by the replicative, four subunit DNA pol α-primase complex within microsatellite sequences is unique for each microsatellite allele examined (18). Strong DNA pol α-primase pause sites within a $[TC]_{20}$ allele were caused by triplex DNA formation between the nascent DNA primer-template duplex and the downstream template. We have observed strong, triplex-related pause sites for termination by pol κ within the same $[TC]_{20}$ allele (data not shown). To explore the generality of DNA structure-induced polymerase pausing within common microsatellites, we examined termination within mononucleotide sequences. We observed intense, polar pause sites within a $[T]_{11}$ allele during DNA synthesis by both pol κ and pol α-primase (Figures 2 and 3). However, we continued to observe the prominent $[T]_{11}$ polar pause sites in the presence of 7-deaza-dATP (Figures 2 and S1). Therefore, the strong pause sites within the $[T]_{11}$ allele do not result from intramolecular triplex DNA formation. We considered whether the conformation of the duplex primer stem formed during DNA synthesis through the microsatellite could affect polymerase termination. Poly(dA/dT) sequences are associated with altered DNA flexibility and the formation of bent DNA (49). A unique feature of bent DNA structure is that the tract displays a polarity, with the minor groove progressively narrowing from the 5′ end to the 3′ end of the track (50,51). DNA bending is also temperature dependent (52). Both of these features, polarity and temperature-dependence, were observed for the pol κ pause sites within the $[T]_{11}$ template (Figures 2 and 4). Moreover, the cooperative unit for bending is ∼5 bp, and regions of uniform bent structure are observed in tracts longer than 7 bp (50,51). The major polymerase pause sites at T3 and T2 would correspond to a primer duplex stem length of eight and nine T–A basepairs, respectively. We propose that the observed pause sites within the $[T]_{11}$ template result from structural changes to the duplex primer-template stem that disrupt polymerase–DNA interactions. Bent DNA displays a gradual shortening of the N3T-N1A hydrogen bond (51), and DNA polymerases utilize hydrogen bonding with minor groove acceptor sites to position the DNA primer stem within the active site for catalysis (53). Alternatively, bent DNA formed within the duplex primer stem may affect polymerase DNA affinity during synthesis. If the bent DNA model is correct, then we must also conjecture that the nucleotide composition of the nascent DNA strand directly affects the biochemical properties of the duplex stem, because we observed less pausing within an $[A]_{11}$ allele in the same sequence context (Table 2). We note that along the $[A]_{11}$ tract, a pause site is observed at position A1, at the very 3′ end of the allele. Possibly, the degree of duplex primer-stem bending is reduced or occurs with a longer phasing for the $A_{template}$–T_{primer} sequence, relative to the $T_{template}$–A_{primer} sequence. Further experimentation is needed to fully test this model.

The rate of replication fork progression is not uniform through the eukaryotic genome (54,55), and replication fork stalling can be observed under non-stressed conditions (14). Previously, we proposed that the strong strand bias for pol α-primase termination within di- and tetranucleotide alleles may lead to a non-uniform rate of lagging strand DNA synthesis. The new results reported here for intense DNA polymerase pausing within poly(dT) sequences demonstrate that the inhibition of DNA synthesis within common microsatellites is sequence-specific. Thus, primary DNA sequence and accompanying DNA structural changes may be a factor contributing to nonrandom replication fork movements observed *in vivo*. In addition, the rate of replication fork progression through microsatellites will be controlled by the identity of the active polymerase, as we have shown that polα-primase termination at the $[T]_{11}$ allele is much more intense than is pol κ termination (Figure 3C).

CONCLUSION

The human genome contains an abundance of poly (dA/dT). Unit-length changes in microsatellite alleles can impact gene expression, and microsatellites are under selective pressure during evolution. The primary mutational mechanism at microsatellite alleles has been shown to be DNA polymerase slippage. We report here a very unusual mutational spectrum of non-canonical base insertion errors produced *in vitro* by DNA polymerases, particularly pol κ, within a $[T]_{11}$ allele. These errors effectively interrupt the microsatellite sequence, and may be a mechanism whereby microsatellites are stabilized in the genome. We have previously reported the production

of both canonical (unit-length) and non-canonical errors within di- and tetranucleotide repeats during *in vitro* DNA synthesis by DNA pol β (11). The non-canonical errors observed include single base deletions or complex base substitution and deletions, all of which produced interrupted alleles from pure microsatellites. Our *in vitro* data for DNA polymerases demonstrate the highly dynamic nature of microsatellite mutagenesis, in that not only the length but also the homogeneity of the microsatellite is in continual flux. In support of the concept of a microsatellite life-cycle (7), we have observed directly the creation of polymorphic, common microsatellite alleles through expansion and contraction errors of unit length. Conversely, as reported here, we have observed the death of common alleles through base interruption errors that create shorter, interrupted alleles.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This study was funded by National Institutes of Health [CA100060] and the Jake Gittlen Cancer Research Foundation. We thank Dr Joann Sweasy, Dr Kataryna Makova, Yogeshwar Kelkar and Sandeep Shah for their critical reading of the manuscript and helpful comments. Funding to pay the Open Access publication charges for this article was provided by the Gittlen Cancer Research Foundation.

Conflict of interest statement. None declared.

REFERENCES

1. Subramanian,S., Mishra,R.K. and Singh,L. (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol.*, **4**, R13.
2. Chambers,G.K. and MacAvoy,E.S. (2000) Microsatellites: consensus and controvery. *Comp. Biochem. Physiol. Part B*, **126**, 455–476.
3. Kashi,Y. and King,D.G. (2006) Simple sequence repeats as advantageous mutators in evolution. *Trends Genet.*, **22**, 253–259.
4. Rothenburg,S., Koch-Nolte,F., Rich,A. and Haag,F. (2001) A polymorphic dinucleotide repeat in the rat nucleolin gene forms Z-DNA and inhibits promoter activity. *Proc. Natl Acad. Sci. USA*, **98**, 8985–8990.
5. Rockman,M.V. and Wray,G.A. (2002) Abundant raw material for cis-regulatory evolution in humans. *Mol. Biol. Evol.*, **19**, 1991–2004.
6. Kruglyak,S., Durrett,R.T., Schug,M.D. and Aquadro,C.F. (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proc. Natl Acad. Sci. USA*, **95**, 10774–10778.
7. Buschiazzo,E. and Gemmell,N.J. (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays*, **28**, 1040–1050.
8. Levinson,G. and Gutman,G.A. (1987) Slipped strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.*, **4**, 203–221.
9. Sia,E.A., Jinks-Robertson,S. and Petes,T.D. (1997) Genetic control of micrsatellite stability. *Mutat. Res.*, **383**, 61–70.
10. Richards,R.I. and Sutherland,G.R. (1994) Simple repeat DNA is not replicated simply. *Nature Genet.*, **6**, 114–116.
11. Eckert,K.A., Mowery,A. and Hile,S.E. (2002) Misalignment-mediated DNA polymerase beta mutations: comparison of microsatellite and frameshift error rates using a forward mutation assay. *Biochem*, **41**, 10490–10498.
12. Li,Y.C., Korol,A.B., Fahima,T., Beiles,A. and Nevo,E. (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.*, **11**, 2453–2465.
13. Bacolla,A. and Wells,R.D. (2004) Non-B DNA conformations, genomic rearrangements, and human disease. *J. Biol. Chem.*, **279**, 47411–47414.
14. Mirkin,E.V. and Mirkin,S.M. (2007) Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.*, **71**, 13–35.
15. Wang,G. and Vasquez,K.M. (2006) Non-B DNA structure-induced genetic instability. *Mutat. Res.*, **598**, 103–119.
16. Wells,R.D. (2007) Non-B DNA conformations, mutagenesis and disease. *Trends Biochem. Sci.*, **32**, 271–278.
17. Sinden,R.R. (1994) *DNA Structure and Function*. Academic Press, New York.
18. Hile,S.E. and Eckert,K.A. (2004) Positive correlation between DNA polymerase alpha-primase pausing and mutagenesis within polypyrimidine/polypurine microsatellite sequences. *J. Mol. Biol.*, **335**, 745–759.
19. Sweasy,J.B., Lauper,J.M. and Eckert,K.A. (2006) DNA polymerases and human diseases. *Radiat. Res.*, **166**, 693–714.
20. Bebenek,K. and Kunkel,T.A. (2004) Functions of DNA polymerases. *Adv. Protein Chem.*, **69**, 137–165.
21. Lehmann,A.R. and Fuchs,R.P. (2006) Gaps and forks in DNA replication: Rediscovering old models. *DNA Repair*, **5**, 1495–1498.
22. Lehmann,A.R., Niimi,A., Ogi,T., Brown,S., Sabbioneda,S., Wing,J.F., Kannouche,P.L. and Green,C.M. (2007) Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair*, **6**, 891–899.
23. Prakash,S., Johnson,R.E. and Prakash,L. (2005) Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu. Rev. Biochem.*, **74**, 317–353.
24. Ogi,T., Kannouche,P. and Lehmann,A.R. (2005) Localisation of human Y-family DNA polymerase kappa: relationship to PCNA foci. *J. Cell. Sci.*, **118**, 129–136.
25. Washington,M.T., Johnson,R.E., Prakash,L. and Prakash,S. (2002) Human DINB1-encoded DNA polymerase kappa is a promiscuous extender of mispaired primer termini. *Proc. Natl Acad. Sci. USA*, **99**, 1910–1914.
26. Wolfe,W.T., Washington,M.T., Prakash,L. and Prakash,S. (2003) Human DNA polymerase kappa uses template-primer misalignment as a novel means for extending mispaired termini and for generating single-base deletions. *Genes Dev.*, **17**, 2191–2199.
27. Garcia-Diaz,M. and Kunkel,T.A. (2006) Mechanism of a genetic glissando: structural biology of indel mutations. *Trends Biochem. Sci.*, **31**, 206–214.
28. Bebenek,K. and Kunkel,T.A. (2000) Streisinger revisited: DNA synthesis errors mediated by substrate misalignments. *Cold Spring Harb. Symp. Quant. Biol.*, **65**, 81–91.
29. Opresko,P.L., Shiman,R. and Eckert,K.A. (2000) Hydrophobic interactions in the hinge domain of DNA polymerase ß are important but not sufficient for maintaining fidelity of DNA synthesis. *Biochemistry*, **39**, 11399–11407.
30. Eckert,K.A. and Yan,G. (2000) Mutational analyses of dinucleotide and tetranucleotide microsatellites in Escherichia coli: influence of sequence on expansion mutagenesis. *Nucleic Acids. Res.*, **28**, 2831–2838.
31. Eckert,K.A., Hile,S.E. and Vargo,P.L. (1997) Development and use of an in vitro HSV-tk forward mutation assay to study eukaryotic DNA polymerase processing of DNA alkyl lesions. *Nucleic Acids. Res.*, **25**, 1450–1457.
32. Toth,G., Gaspari,Z. and Jurka,J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.*, **10**, 967–981.
33. Jacob,K.D. and Eckert,K.A. (2007) Escherichia coli DNA polymerase IV contributes to spontaneous mutagenesis at coding sequences but not microsatellite alleles. *Mutat. Res.*, **619**, 93–103.
34. Tran,H.T., Keen,J.D., Kricker,M., Resnick,M.A. and Gordenin,D.A. (1997) Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Mol. Cell. Biol.*, **17**, 2859–2865.
35. Sia,E.A., Kokoska,R.J., Dominska,M., Greenwell,P. and Petes,T.D. (1997) Microsatellite instability in yeast: dependence on repeat unit size and mismatch repair genes. *Mol. Cell. Biol.*, **17**, 2851–2858.

36. Wagner,L.A., Weiss,R.B., Driscoll,R., Dunn,D.S. and Gesteland,R.F. (1990) Transcriptional slippage occurs during elongation at runs of adenine or thymine in Escherichia coli. *Nucleic Acids Res.*, **18**, 3529–3535.
37. Sen,A. and Graslund,A. (2000) Structural constraints regulating triple helix formation by A-tracts. *Biophys. Chem.*, **88**, 69–80.
38. Johnson,R.E., Prakash,L. and Prakash,S. (2005) Biochemical evidence for the requirement of Hoogsteen base pairing for replication by human DNA polymerase iota. *Proc. Natl Acad. Sci. USA*, **102**, 10466–10471.
39. Eckert,K.A. and Kunkel,T.A. (1993) Fidelity of DNA synthesis catalyzed by human DNA polymerase alpha and HIV-1 reverse transcriptase: Effect of reaction pH. *Nucleic Acids Res.*, **21**, 5212–5220.
40. Taylor,J.S., Durkin,J.M. and Breden,F. (1999) The death of a microsatellite: a phylogenetic perspective on microsatellite interruptions. *Mol. Biol. Evol.*, **16**, 567–572.
41. Blanquer-Maumont,A. and Crouau-Roy,B. (1995) Polymorphism, monomorphism, and sequences in conserved microsatellites in primate species. *J. Mol. Evol.*, **41**, 492–497.
42. Petes,T.D., Greenwell,P.W. and Dominska,M. (1997) Stabilization of microsatellite sequences by variant repeats in the yeast Saccharomyces cerevisiae. *Genetics*, **146**, 491–498.
43. Zhong,X., Garg,P., Stith,C.M., McElhinny,S.A., Kissling,G.E., Burgers,P.M. and Kunkel,T.A. (2006) The fidelity of DNA synthesis by yeast DNA polymerase zeta alone and with accessory proteins. *Nucleic Acids Res.*, **34**, 4731–4742.
44. Bebenek,K., Garcia-Diaz,M., Blanco,L. and Kunkel,T.A. (2003) The frameshift infidelity of human DNA polymerase lambda. Implications for function. *J. Biol. Chem.*, **278**, 34685–34690.
45. Matsuda,T., Bebenek,K., Masutani,C., Rogozin,I.B., Hanaoka,F. and Kunkel,T.A. (2001) Error rate and specificity of human and murine DNA polymerase eta. *J. Mol. Biol.*, **312**, 335–346.
46. Eckert,K.A. and Hile,S.E. (1998) Alkylation-induced frameshift mutagenesis during in vitro DNA synthesis by DNA polymerases a and ß. *Mutat. Res.*, **422**, 255–269.
47. Ohashi,E., Bebenek,K., Matsuda,T., Feaver,W.J., Gerlach,V.L., Friedberg,E.C., Ohmori,H. and Kunkel,T.A. (2000) Fidelity and processivity of DNA synthesis by DNA polymerase kappa, the product of the human DinB1 gene. *J. Biol. Chem.*, **275**, 39678–39684.
48. Zhang,Y., Yuan,F., Xin,H., Wu,X., Rajpal,D.K., Yang,D. and Wang,Z. (2000) Human DNA polymerase kappa synthesizes DNA with extraordinarily low fidelity. *Nucleic Acids Res.*, **28**, 4147–4156.
49. Neidle,S. (1999) *Oxford Handbook of Nucelic Acid Structure*. Oxford University Press, New York.
50. Crothers,D.M., Haran,T.E. and Nadeau,J.G. (1990) Intrinsically bent DNA. *J. Biol. Chem.*, **265**, 7093–7096.
51. Nadeau,J.G. and Crothers,D.M. (1989) Structural basis for DNA bending. *Proc. Natl Acad. Sci. USA*, **86**, 2622–2626.
52. Koo,H.S., Wu,H.M. and Crothers,D.M. (1986) DNA bending at adenine. thymine tracts. *Nature*, **320**, 501–506.
53. Kiefer,J.R., Mao,C., Braman,J.C. and Beese,L.S. (1998) Visualizing DNA replication in a catalytically active Bacillus DNA polymerase crystal. *Nature*, **391**, 304–307.
54. Raghuraman,M.K., Winzeler,E.A., Collingwood,D., Hunt,S., Wodicka,L., Conway,A., Lockhart,D.J., Davis,R.W., Brewer,B.J. and Fangman,W.L. (2001) Replication dynamics of the yeast genome. *Science*, **294**, 115–121.
55. Cha,R.S. and Kleckner,N. (2002) ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science*, **297**, 602–606.