BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

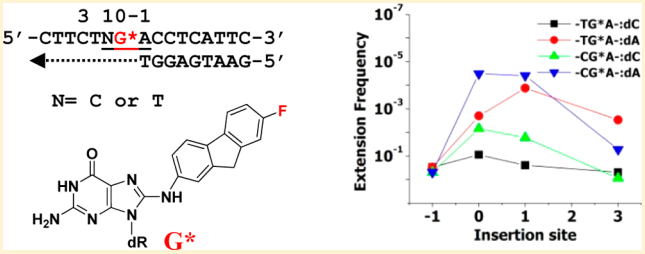
Sequence Effects on Translesion Synthesis of an Aminofluorene–DNA Adduct: Conformational, Thermodynamic, and Primer Extension Kinetic Studies

V. G. Vaidyanathan and Bongsup P. Cho*

Department of Biomedical and Pharmaceutical Sciences, College of Pharmacy, University of Rhode Island, Kingston, Rhode Island 02881, United States

Supporting Information

ABSTRACT: The DNA sequence effect is an important structural factor for determining the extent and nature of carcinogen-induced mutational and repair outcomes. In this study, we used two 16-mer template sequences, TG*A [d(5′-CTTCTTG*ACCTCATTC-3′)] and CG*A [d(5′-CTTCTCG*ACCTCATTC-3′)], to study the impact of the 5′-flanking nucleotide (T vs C) on aminofluorene (AF)-induced stacked (S)/major groove (B)/wedge (W) conformational heterogeneity during a simulated translesion synthesis. In addition, we probed the sequence effect on nucleotide insertion efficiencies catalyzed by the Klenow fragment (exonuclease-deficient) of DNA polymerase I. Our ^{19}F NMR/ICD/DSC results showed that AF in the CG*A duplex sequence adopts a greater population of S-conformer than the TG*A sequence. We found that the S conformer of CG*A thermodynamically favors insertion of A over C at the lesion site (n). Significant stalling occurred at both the prelesion ($n - 1$) and lesion (n) sites; however, the effect was more persistent for the S conformer of CG*A than TG*A at the lesion site (n). Kinetics show that relative nucleotide insertion frequencies (f_{ins}) were greater for TG*A than the S conformer of CG*A for either dCTP or dATP at the lesion site (n), and the insertion rate was significantly reduced at immediate upstream base pairs (n, $n + 1$). Taken together, the results provide insight into how the mutagenic AF could exhibit an S/B/W equilibrium in the active site of a polymerase, causing different mutations. This work represents a novel structure–function relationship in which adduct structure is directly linked to nucleotide insertion efficiency in a conformation-specific manner during translesion DNA synthesis.



The human genome is under constant attack by various endogenous and exogenous chemical and radiation assaults, producing different types of lesions.[1] The majority of DNA damage is repaired by various mechanisms.[2,3] However, unrepaired lesions can stall replication, which in turn recruits specialized translesional DNA polymerases to bypass the lesion. Unrepaired DNA lesions are often mutagenic because of their aberrant geometry and structure that alters their templating capacity, hence initiating cancer.[2] The nature of mutation is inherently complex, but it is thought to be influenced by various factors, including (i) the structure or conformation of the lesion in the confines of the DNA polymerase active site, (ii) the neighboring DNA sequence context in which the lesion is embedded, (iii) the identity of the incoming nucleotide opposite the lesion, and (iv) the specific polymerase(s) involved in translesion DNA synthesis (TLS).[2,4,5]

Arylamines are notorious environmental mutagens implicated in the etiology of various sporadic human cancers.[6−8] Aminofluorene is a prototype arylamine carcinogen and produces two major C8-substituted dG adducts, AAF and AF, in vivo, depending on the presence and absence of an acetyl group on the central nitrogen, respectively (Figure 1a).[9,10] Despite their structural differences, the two adducts adopt three distinctive conformational motifs, depending on the location of the hydrophobic aminofluorene moiety: the major groove-binding B-type (B), the base-displaced stacked (S), and the minor groove-binding wedge (W) conformers (Figure 1).[11−14] Previously, the arylamine-induced S/B/W heterogeneity has been found to be modulated by the surrounding sequence context.[15,16] While AF produces mostly point mutations, AAF strongly blocks replicative polymerases, producing both point and frameshift mutations. These mutagenic differences between the two adducts could be related to the fact that AF in fully paired duplexes exists in an equilibrium of S/B conformers, whereas N-acetylated AAF adopts a more complex mixture of S/B/W conformers. The S conformer of AF in the coding position of a DNA polymerase active site could prevent critical polymerase conformational changes during dNTP incorporation that promote polymerase arrest. With the B conformation, however, the hydrophobic aminofluorene moiety could be accommodated in the major groove with minimal distortion in the modified templating guanine, which forms Watson–Crick

Received: November 24, 2011
Revised: December 31, 2011
Published: February 13, 2012

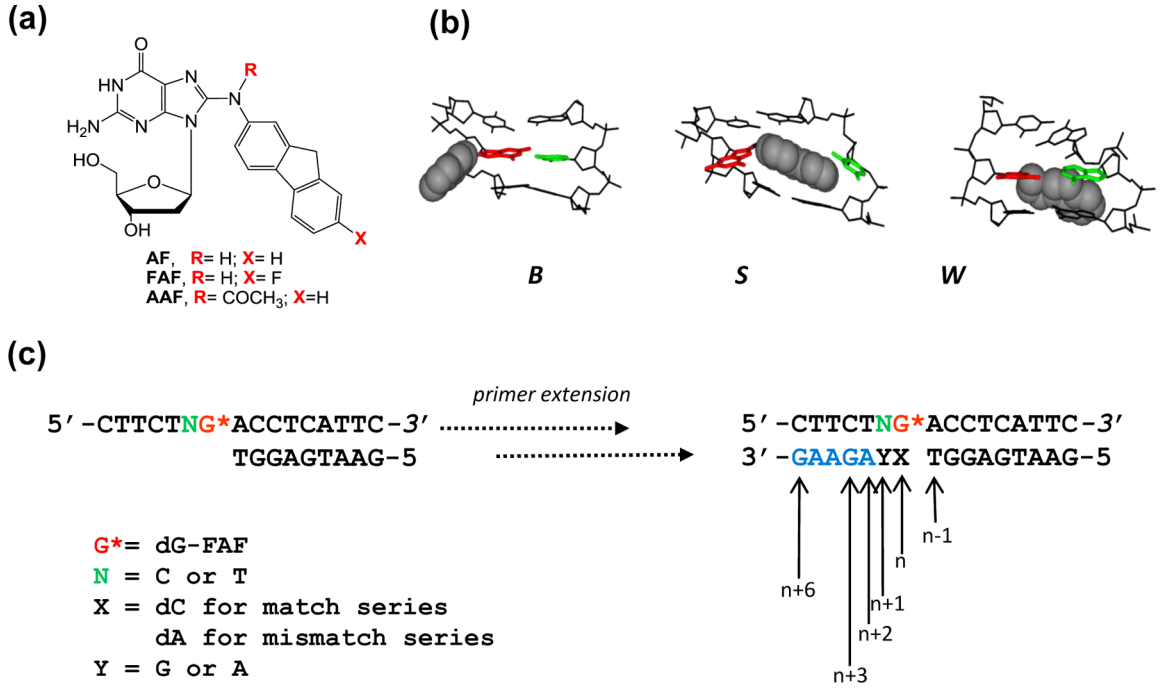


Figure 1. (a) Structure of the FAF adduct, *N*-(2′-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene. (b) Molecular models of different conformations of AF-adducted DNA major groove-binding (B), base-displaced (S), and minor groove-binding (W) conformers. (c) Primer–template sequences in a simulated translesion synthesis.

base pairs with the complementary incoming cytosine. While AAF requires a specialized polymerase to bypass this lesion, AF can be bypassed by a replicative polymerase, albeit with low processivity. It is worth noting that crystal structure studies of T7 DNA polymerase in a complex with an AF-modified template primer sequence failed to reveal clear electron density for the adduct at the templating position, which could be due to the existence of equilibrating S/B conformers.[17] A shift in equilibrium between different conformations was found to depend on the nature of the bases flanking and opposite the lesion.[16,18,19]

There are numerous examples in which site-specifically modified bulky lesions exhibit different mutation and repair outcomes depending on the nature of the flanking base sequences. Studies with AF embedded in a replicating plasmid vector in simian kidney (COS-7) cells have shown an unusually high mutation frequency (70%) with guanine 5′ to the lesion and only 2–4% with other bases positioned 5′ to the lesion site.[5] Loechler et al.[20] have shown that 5′-TG* sequences consistently produce a higher fraction of G → T mutations for the major BaP–N^2-dG adduct in comparison with 5′-CG*, 5′-GG*, or 5′-AG* sequences. We have shown previously that purine bases (G > A ≫ C > T) 3′-flanking the AF lesion promote the S conformation with greater efficiency in the UvrABC nucleotide excision repair (NER) system in *Escherichia coli*. The molecular basis for this novel "conformation-specific NER" appears to be disruption of Watson–Crick bonds at the lesion site.[16,21] Zou et al.[22] have shown that AF and AAF adducts in the TG*T sequence context are incised more efficiently in *E. coli* by a factor of 1.7 than in the CG*C context. It was argued that the T:A flanking base pairs in the former allow for local bending and flexibility, thus promoting greater conformational heterogeneity.[22] Similar "flanking T effects" have been reported for the *trans-anti*-BP–dG adduct.[18,23,24] Primer extension assays on AF by high-fidelity polymerases revealed that dCTP was preferentially incorporated opposite the lesion, followed by insertion of an incorrect dAMP opposite the lesion, which leads to a G → T transversion.[5] Steady-state kinetic studies with Klenow fragment have shown significant reductions in incorporation rates around and downstream of the lesion.[25] A recent study has demonstrated that the translesional synthesis of AF in Dpo4 leads to full-length product as well as a semitargeted mutation.[26] Assays of the truncated form of pol κ (pol κΔC) with AF resulted in one- and two-base deletions with high levels of misincorporation.[27] While the AF lesion has been studied extensively, little is known about the sequence dependence of its conformational heterogeneity, thermodynamics, and polymerase-mediated extension assay in TLS.

We previously conducted a systematic calorimetric (DSC) and spectroscopic [^{19}F NMR/induced circular dichroism (ICD)/UV] investigation of a simulated TLS model system, in which a 16-mer template containing FAF was annealed with growing primers in both dC-match and dA-mismatch series [TG*A series (Figure 1c)].[28,29] In this work, we studied essentially the same TLS system except that the 5′-flanking dT of the template was replaced with dC [CG*A series (Figure 1c)]. Additionally, we performed steady-state primer extension kinetics experiments with Klenow fragment exonuclease-deficient (Kf-exo$^-$) activity on both TG*A and CG*A TLS systems and conducted comparative analysis. The results provide valuable insight into the 5′-flanking sequence effect of the AF-induced S/B/W conformational heterogeneity in a simulated TLS system.

EXPERIMENTAL PROCEDURES

Materials and Methods. Crude oligodeoxynucleotides (10–15 μmol) in desalted form were purchased from Operon (Eurofin, Huntsville, AL) and purified by reverse phase high-performance liquid chromatography (RP-HPLC). All HPLC

dx.doi.org/10.1021/bi2017443 | *Biochemistry* 2012, 51, 1983–1995

solvents were purchased from Fisher Inc. (Pittsburgh, PA). The HPLC system consisted of a Hitachi EZChrom Elite HPLC system with an L2450 diode array as a detector and a Clarity column (10 mm × 150 mm, 5 μm) (Phenomenex, Torrance, CA). The mobile phase system involved a 30 min linear gradient profile from 3 to 12% (v/v) acetonitrile with 100 mM ammonium acetate buffer (pH 7.0) at a flow rate of 3.0 mL/min. T4 polynucleotide kinase (T4 PNK) was purchased from USB Chemicals. [γ-^{32}P]ATP (specific activity of 6000 Ci/mmol) was purchased from Perkin-Elmer Radiochemicals (Boston, MA). Kf-exo$^-$ (D424A) was received as a gift from C. Joyce (Yale University, New Haven, CT).

Preparation of an FAF-Modified Template. An FAF-modified 16-mer CG*A DNA template, d(5′-CTTCTCG* ACCTCATTC-3′, G* = FAF), was prepared according to a published procedure.[28] Briefly, an unmodified 16-mer oligonucleotide was treated with *N*-acetoxy-*N*-trifluoroacetyl-7-fluoro-2-aminofluorene, an activated derivative of 7-fluoro-2-aminofluorene. The reaction was monitored over a period of 24 h by RP-HPLC using the linear gradient described above. The modified strand was purified by RP-HPLC and characterized by mass spectrometry as described previously.[28,29]

NMR. Approximately 30 optical density (OD) units of pure modified oligonucleotides was annealed with primer sequences to produce appropriate template−primer sequences (see Figure 1c for sequences). The samples were filtered by ultracentrifugation using a Pall Microsep MF centrifugal device (Yellow, M_r cutoff of 1000). The centrifuged samples were dissolved in 300 μL of a neutral buffer [10% D_2O/90% H_2O mixture containing 100 mM NaCl, 10 mM sodium phosphate (pH 7.0), and 100 μM tetrasodium EDTA], filtered through a 0.2 μm membrane filter, and placed in a Shigemi tube for NMR experiments.

All ^{1}H and ^{1}H-decoupled ^{19}F NMR results were recorded using a dedicated 5 mm ^{19}F/^{1}H dual probe on a Bruker DPX400 Avance spectrometer operating at 400.0 and 376.5 MHz, respectively. Imino proton spectra were obtained using phase sensitive jump−return sequences at 5 °C and referenced to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). ^{19}F NMR spectra were referenced to $CFCl_3$ by assigning external hexafluorobenzene in C_6D_6 at −164.90 ppm. One-dimensional ^{19}F NMR spectra at 5−60 °C were obtained by collecting 65536 points using a 37664 Hz sweep width and a recycle delay of 1.0 s between acquisitions. A total of 1200 scans were acquired for each spectrum. The spectra were processed by zero-filling, exponential multiplication using 20 Hz line broadening, and Fourier transformation. The peak areas were baseline corrected and integrated using XWIN NMR (Bruker, Billerica, MA). Two-dimensional exchange (EXSY) ^{19}F NMR spectra were recorded in phase-sensitive mode using a NOESY pulse sequence: sweep width of 4529 Hz, 1024 complex data points in t_2, 256 complex free induction decays in t_1, 96 scans, 16 dummy scans, recycle delays of 1.0 s, and mixing time of 400 ms. The data were subjected to sine bell apodization using 2 Hz line broadening in both dimensions and then zero-filled before Fourier transformation of the 1024 × 256 data matrix. The data were not symmetrized. Complete line shape analysis was conducted using WINDNMR-Pro (version 7.1.6, *Journal of Chemical Education* Software Series, Reich, H. J., University of Wisconsin, Madison, WI).

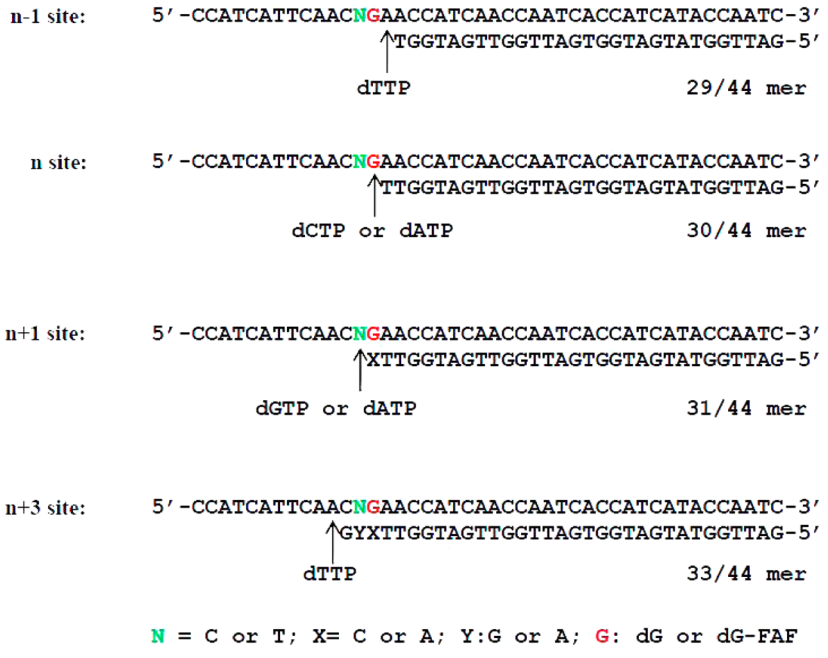
Circular Dichroism (CD). CD measurements were taken on a Jasco J-810 spectropolarimeter equipped with a variable Peltier temperature controller. Typically, 2 OD of a FAF-modified template strand was annealed with 1 equiv of primer strands. The samples were dissolved in 400 μL of a neutral buffer [0.2 M NaCl, 10 mM sodium phosphate (pH 7.0), and 0.2 mM EDTA] and placed in a 1 mm path length cell. The sample was heated at 85 °C for 5 min and then cooled to 15 °C, over a 10 min period to ensure duplex formation. Spectra were recorded from 200 to 400 nm at a rate of 50 nm/min; the final data were averaged from 10 accumulations. Data points were acquired every 0.2 nm with a 2 s response time.

Differential Scanning Calorimetry (DSC). DSC measurements were performed using a Nano-DSC calorimeter from TA Instruments. The experiments were conducted using a previously reported procedure with a slight modification in data processing.[28] Briefly, samples (1 mL) were degassed for 15 min at 25 °C by using Thermovac sample degassing and a thermostatic pump with constant stirring. Solutions were loaded into sample and reference cells with a 2.5 mL micropipet. Samples were carefully pipetted into the cells to minimize air bubbles. After both cells had been filled, they were capped, and an external pressure (~3 atm) was applied to prevent evaporation of the sample solution. Raw data were collected as microwatts versus temperature. Template−primer solutions were prepared by dissolving desalted samples in a pH 7.0 buffer solution consisting of 20 mM sodium phosphate and 0.1 M NaCl. A 0.1 mM template−primer solution was scanned against buffer from 15 to 85 °C at a rate of 0.75 °C/min, which was repeated five times. The first scan of each sample run was discarded for analysis because of the deviation in analysis. A buffer versus buffer scan was used as a control and subtracted from the sample scan and normalized for heating rate. This results in base-corrected ΔC_p^{ex} versus temperature curves. Each transition shows negligible changes in the heat capacities between the pre- and post-transition states, thus assumed to be zero. Pre- and post-transition temperatures were marked and used for determining enthalpy (ΔH). Other thermodynamic parameters were calculated by using the previously reported procedure.[28]

Primer Extension Assay. The primer (9-mer) used in the NMR/DSC studies was 5′-radiolabeled using [γ-^{32}P]ATP and T4 polynucleotide kinase (T4 PNK) by following the manufacturer's protocol. The 5′ ^{32}P-labeled primer (100 pmol) was annealed to either an unmodified or adducted template oligonucleotide (120 pmol) by being heated to 95 °C for 5 min and then slowly cooled to room temperature. The primer−template sequence (100 nM) was incubated with varying concentrations of Kf-exo$^-$ (0−100 nM) for 5 min to form a binary complex. The reaction was initiated by adding a dNTP/$MgCl_2$ (each dNTP at 250 μM with 5 mM $MgCl_2$) solution to a binary mixture, the mixture incubated at 20 °C for 10 min in Tris buffer [50 mM Tris-HCl (pH 7.4), 50 μg/mL BSA, and 5% (v/v) glycerol], and the reaction quenched with 50 μL of a 50 mM EDTA (pH 8.0)/95% formamide solution. The quenched sample was heated to 95 °C for 5 min and immediately cooled on ice. The products were resolved with a denaturing polyacrylamide gel [20% polyacryamide (w/v)/7 M urea] and electrophoresed at 2000 V for 4 h. The gel was exposed on a Kodak phosphor imaging screen overnight and scanned with a Typhoon 9410 variable mode imager. The band intensities were quantitated using ImageQuantTL from GE Healthcare.

For translesion bypass DNA synthesis, a 25-mer primer was used with the modified dG at position 31 of the 44-mer template as illustrated in Figure 5. For the standing-start

Scheme 1. Oligonucleotide Sequences Used for Steady-State Kinetic Assays



kinetic assays, longer primers (29−33-mers) were used to examine insertion or extension efficiencies in the vicinity of the lesion as shown in Scheme 1. The enzyme concentration (2.5 or 5.0 nM) was kept constant and the reaction time varied so that product formation (i.e., velocities) was linear.

Determination of Kinetic Parameters of Nucleotide Incorporation. Steady-state kinetic parameters for incorporation of the nucleotide opposite the unmodified and AF-modified templates were studied by following the procedures reported previously.[25] The reactions were performed at 20 °C. For the unmodified sequence, reactions were performed for ≤1 min for correct nucleotide incorporation and for up to 45 min for incorrect nucleotide incorporation. The percentage of primer extended in kinetic assays was determined by taking the ratio of extended primer to the total amount of primer (unextended + extended primer). Initial velocities were calculated by plotting product concentration versus time. The kinetic parameters k_{cat} and K_m were determined by a non-linear least-squares fit of the initial velocity versus dNTP concentration using SigmaPlot 11.0 (Systat Software Inc.). The efficiency of nucleotide insertion by Kf-exo$^-$ was calculated as k_{cat}/K_m. The relative insertion (f_{ins}) or extension (f_{ext}) efficiencies were calculated as $(k_{cat}/K_m)_{mismatch\ or\ lesion}/(k_{cat}/K_m)_{unmodified\ control}$.

RESULTS

Model System. Figure 1 shows the chemical structures of AF, FAF (a ^{19}F-tagged model for AF), and AAF and the sequences for the simulated TLS systems used in previous studies (TG*A, G* = dG-FAF) and this study (CG*A, G* = dG-FAF). Figure 1b depicts the three FAF-induced conformational motifs: B, S, and W.

Five model template−primer sequences (Figure 1c) were prepared by annealing the FAF-modified 16-mer strand with primers of the appropriate length (n − 1, n, n + 1, n + 3, and n + 6, where n is the lesion site). For each sequence, two distinct TLS systems were created, depending on the nature of the base added opposite the lesion site: dC-match (X = dC) and dA-mismatch (X = dA) series. Preparation, purification, and ESI-MS characterization of the FAF-modified CG*A 16-mer DNA template were conducted according to the general procedure described previously.[28] Basically, we studied two template sequences, TG*A [d(5′-CTTCTTG*ACCTCATTC-3′)] and CG*A [d(5′-CTTCTCG*ACCTCATTC-3′)], which differ from one another by the presence of either a T or C positioned 5′ to FAF.

Dynamic ^{19}F NMR. Figure 2 shows the dynamic (5−50 °C) ^{19}F NMR spectra of the CG*A TLS system at various primer positions for the dC-match and dA-mismatch series. The n + 6 duplexes in the dC-match series, which represent a fully paired end point in the TLS process, displayed clearly defined S/B conformational heterogeneity, particularly at low temperatures. The ^{19}F signal assignments were made on the basis of H/D isotope, exchange spectroscopy (EXSY), and $ICD_{290-350}$ studies, according to the general protocols established previously for the TG*A TLS system and other FAF-modified duplexes.[29]

The ^{19}F NMR characteristics of the CG*A TLS system at the lesion site (n) and the immediate flanking positions (n − 1 and n + 1) were very similar to those of the TG*A TLS system.[29] Specifically, the lesions in these primer positions in the dC-match series were characterized by single broad ^{19}F signals in a narrow range of −117.0 to −118 ppm (Figure 2a) with a positive $ICD_{290-350}$ (Figure S4 of the Supporting Information). These data imply an equilibrium of the syn-type S* and anti-type B* conformations with preference for the former. The dA-mismatch series exhibited a similar ^{19}F NMR pattern with a trapezoidal $ICD_{290-350}$, indicating a preference for the syn-type S*/W* conformations. The asterisked conformation indicates a flexible lesion located at the single and double strand junction.

At n + 3, the lesion in the dC-match series exhibited two well-resolved ^{19}F signals at −117.6 and −118.4 ppm at 5 °C (Figure 2a). These signals exhibited H/D isotope effects of 0.23 and 0.03 ppm, respectively, and hence were assigned to the B and S conformers, respectively.[29] The two signals exchanged slowly in the exchange spectroscopy (EXSY) spectra (Figure 2 inset n + 6, 27 °C), coalesced at ∼35 °C, and started to melt at 45 °C. The S/B heterogeneity was largely intact at n + 6, with

dx.doi.org/10.1021/bi2017443 | Biochemistry 2012, 51, 1983−1995

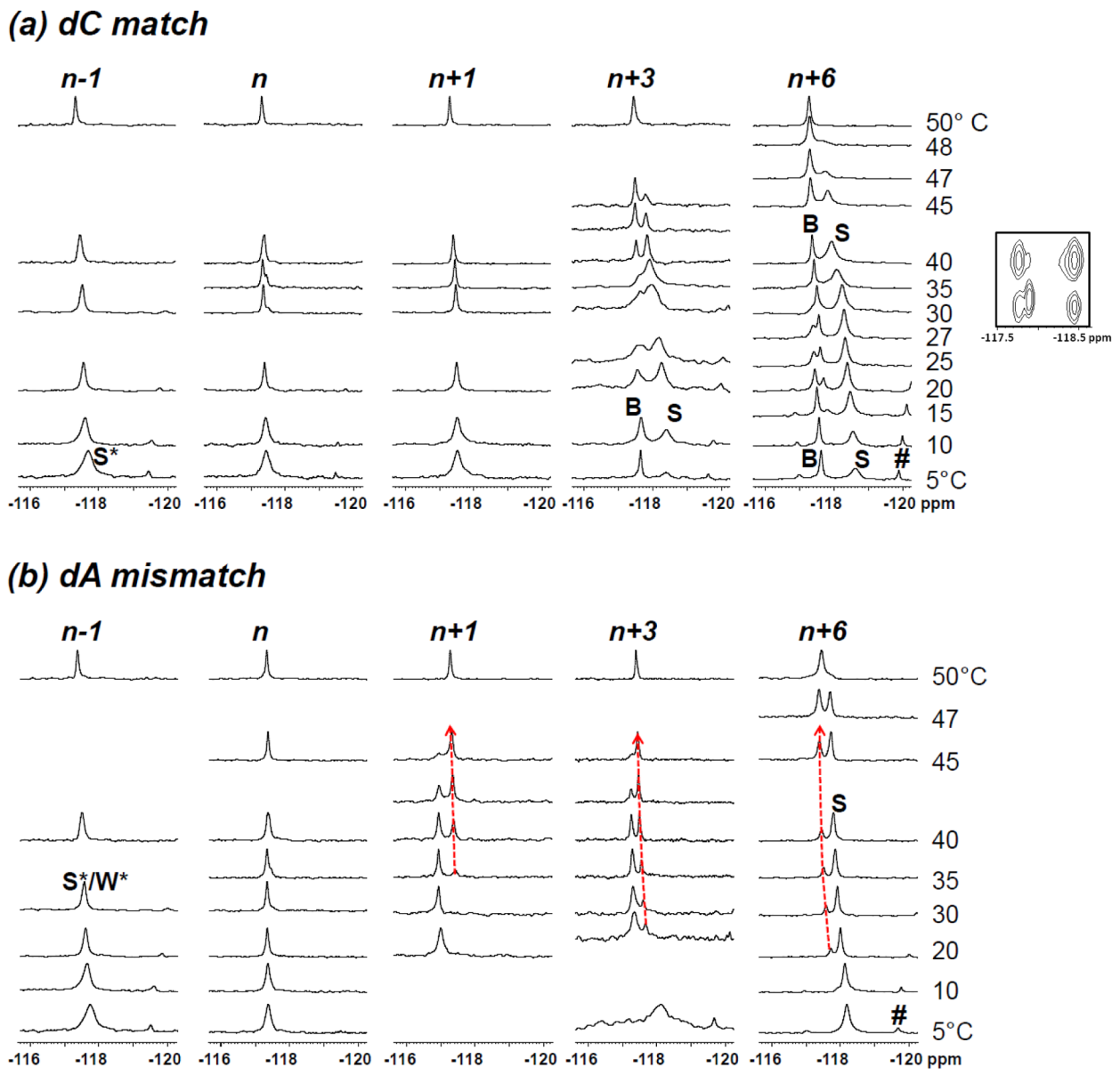


Figure 2. ^{19}F NMR spectra of the CG*A template in a simulated translesion synthesis with (a) dC-match and (b) dA-mismatch series. The red arrow indicates the melting of the duplex. The inset shows the two-dimensional EXSY contour plot of the full duplex of CG*A at 27 °C. S* and W* represent a flexible lesion located at the single and double strand junction with dominant S and W conformations, respectively. # represents an impurity peak.

two well-resolved peaks at −117.5 and −118.6 ppm. Unlike the n + 6 TG*A duplex,[29] however, as the temperature increased the downfield B conformer signal was split into two at −117.4 and −117.6 ppm at 20 °C and coalesced at 30 °C. The two major signals then followed a typical two-site chemical exchange, finally yielding a single sharp signal at 50 °C. As expected, the single upfield S conformer signal exhibited weak H/D shielding (0.07 ppm) with an increase in the D_2O content from 10 to 100%, whereas the split B conformer signals at 0.26 and 0.17 ppm showed H/D effects, indicating the exposed nature of these signals (data not shown). Line shape simulations (Figure S3 of the Supporting Information) indicated that the percent populations of the B and S conformers for the n + 6 fully paired CG*A duplex at 10 °C were 42 and 58%, respectively.

Dynamic ^{19}F NMR profiles (Figure 2b) for the CG*A dA-mismatch series were similar to those of the TG*A system. As such, the *syn*-S* and *syn*-W* signals were in equilibrium with the duplex melting signals (~117.5 ppm, marked as red arrows) throughout primer extension (n + 1 to n + 6).[29] As expected, the chemical shift is similar to that of a denatured FAF-modified template (i.e., 50 °C), regardless of the temperature or the extent of primer extension. The n + 6 dA mismatch duplex exhibited a single prominent signal at −118.2 ppm at 5 °C with a 0.17 ppm H/D shielding when the D_2O content was increased from 10 to 100%. Duplex melting (red arrows) started as low as 20 °C, and the two peaks were fairly sharp up until 47 °C. These results indicated a stability of the dA-mismatch duplex, in which the lesion is located in the minor groove area.[18] A coalescence peak was obtained at 50 °C, which indicated local denaturation at the lesion site leading to the exchange of two conformers.

$ICD_{290-350}$. Figure S4 of the Supporting Information shows the FAF-induced $ICD_{290-350}$ characteristics at various primer elongation steps for the CG*A TLS system. $ICD_{290-350}$ has been shown to be a useful spectral marker for probing FAF-induced

dx.doi.org/10.1021/bi2017443 | *Biochemistry* 2012, 51, 1983−1995

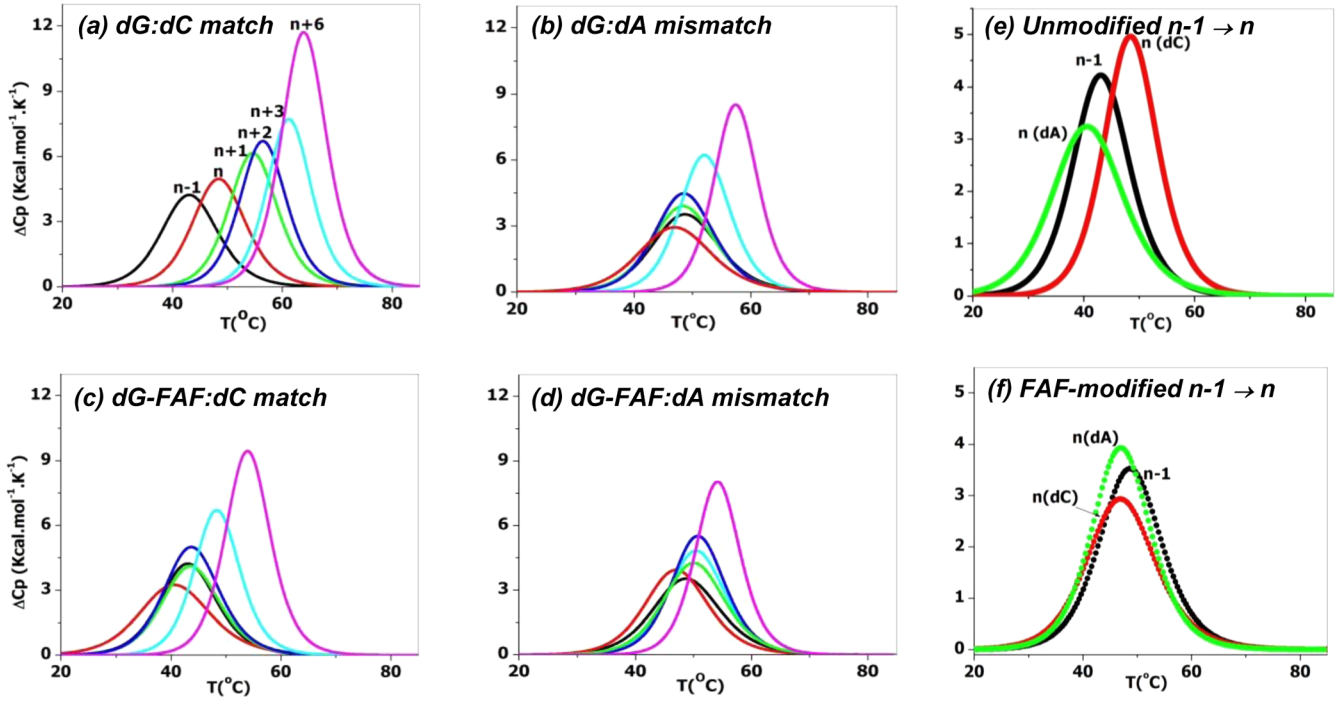


Figure 3. Differential scanning calorimetry (DSC) curves in 20 mM phosphate buffer containing 0.1 M NaCl (pH 7). The total heat of the helix−coil transition of (a) dG:dC-match, (b) dG:dA, (c) dG-FAF:dC-match, and (d) dG-FAF:dA-mismatch template primer series was measured, and the excess heat capacity, ΔC_p^{ex}, is plotted vs temperature. (e and f) DSC curves of an $n - 1$ (black) → n transition with insertion of dC (red) and dA (green) opposite (e) unmodified and (f) FAF-modified G at the primer terminal junction (n).

S/B/W heterogeneity: positive for *syn*-S/W conformers and negative for the *anti*-B conformer.[29,30] In the dC-match series, the ellipticity in the 290−350 nm range decreased as the primer length increased and ended with a slight negative dip. These results indicate a transition from the mostly S*-type conformer at $n - 1$ to $n + 1$ to an S/B mixture at $n + 3$ to $n + 6$. However, the dA-mismatch model remained in the W conformer, as evidenced by its mostly positive trapezoidal ellipticity in the same region. These CD spectral data are consistent with the conformational heterogeneity observed in the ^{19}F NMR spectra described above.

Differential Scanning Calorimetry (DSC). Figure 3 shows plots of calorimetric heat capacity changes (ΔC_p) as a function of temperature for the unmodified (control) and modified CG*A TLS system. The thermodynamic parameters derived from these DSC data are summarized in Table 1.

dC-Match Series. The thermal (T_m) and thermodynamic ($-\Delta H°$ and $-\Delta G_{37\,°C}$) values of the dC-match control TLS series increased incrementally with primer extension (Table 1). Insertion of dC at the primer terminus (n) resulted in increases in $-\Delta H$ (6.3 kcal/mol), $-\Delta G_{37\,°C}$ (0.8 kcal/mol), and ΔT_m (3.5 °C) (Table 2), and this trend continued throughout the entire TLS to $n + 6$, in which all of the pairs in the duplex are supposedly comprised of Watson−Crick base pairs. In contrast, insertion of dC opposite FAF at the primer terminus (n) resulted in decreases in $-\Delta H$ (−5.9 kcal/mol), $-\Delta G_{37\,°C}$ (−0.6 kcal/mol), and ΔT_m (−2.9 °C) (Table 3). The lesion effect was weaker with incorporation of dG at $n + 1$ [$-\Delta H_{n+1} - (-\Delta H_{n-1}) = 2.6$ kcal/mol; $-\Delta G_{37\,°C} \sim 0$ kcal/mol]. The lesion induced a negative thermodynamic impact that persisted until $n + 2$, and then a normal trend resumed at $n + 3$ (Table 1).

dA-Mismatch Series. Addition of dA at the primer terminus (n) of the dA-mismatch TLS control series resulted in major reductions in $-\Delta H$ and ΔT_m by −7.5 kcal/mol and −3.9 °C, respectively (Table 2). Although Watson−Crick base pairs followed the mismatch at n, $-\Delta H$ for the $n - 1$ to $n + 2$ extension was increased by only 4.9 kcal/mol. This is in contrast to the value (17.5 kcal/mol) observed with the aforementioned dC-match control series (Table 1). The Gibbs free energy ($-\Delta G_{37\,°C}$) from the $n - 1$ to $n + 2$ extension was not changed; however, the normal trend resumed at $n + 3$ (Table 1). Unlike the control series, incorporation of dA opposite FAF at lesion site n resulted in a modest increase in enthalpy ($-\Delta H$ = 3.3 kcal/mol) (Table 3). This positive influence persisted until $n + 2$, with the effect especially significant ($-\Delta H$ = 18.8 kcal/mol) at $n + 2$. As usual, a normal trend resumed at $n + 3$.

Table 1. Thermodynamic Parameters Derived from DSC for a Simulated AF-Induced Translesion DNA Synthesis

sequence	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G_{37\,°C}$ (kcal/mol)	T_m (°C)
$n - 1$	53.9 (58.0)[a]	147.7 (162.7)	8.1 (7.5)	48.9 (44.6)
		dC-Match Series		
n	48.0 (64.3)	130.7 (180.6)	7.5 (8.3)	46.0 (48.1)
$n + 1$	56.5 (72.5)	156.3 (201.6)	8.1 (10.0)	48.2 (54.5)
$n + 2$	60.6 (75.5)	168.9 (209.6)	8.3 (10.5)	48.3 (56.1)
$n + 3$	72.2 (82.6)	202.6 (228.0)	9.4 (11.9)	51.8 (60.3)
$n + 6$	85.9 (106.0)	240.5 (295.4)	11.4 (14.4)	57.2 (63.3)
		dA-Mismatch Series		
n	57.2 (50.5)	158.7 (141.3)	8.0 (6.7)	47.3 (40.7)
$n + 1$	57.9 (58.6)	159.4 (165.1)	8.5 (7.4)	50.1 (43.8)
$n + 2$	72.7 (62.9)	205.3 (178.8)	9.0 (7.5)	49.8 (43.9)
$n + 3$	66.2 (74.2)	184.6 (211.2)	8.9 (8.8)	50.9 (48.4)
$n + 6$	82.8 (90.6)	233.1 (258.4)	10.6 (10.5)	54.5 (52.8)

[a]Values of the unmodified control duplex are given in parentheses. Standard deviations for $-\Delta H°$, $-\Delta S°$, $-\Delta G_{37\,°C}$, and T_m are ±3 kcal/mol, ±7 eu, ±0.4 kcal/mol, and ±0.4 °C, respectively. The concentration of oligonucleotides used in these calculations was 0.1 mM.

dx.doi.org/10.1021/bi2017443 | *Biochemistry* 2012, 51, 1983−1995

Table 2

Primer Extension at the Primer Terminus (n) with No Lesion				
unmodified -CGA-	$-\Delta H^a$ (kcal/mol)	$-\Delta G_{37\,°C}{}^b$ (kcal/mol)	$-\Delta S^c$ (eu)	$\Delta T_m{}^d$ (°C)
dC match	6.3 (7.6)[e]	0.8 (1.6)	17.9 (19.8)	3.5
dA mismatch	−7.5 (−1.5)	−0.8 (−0.1)	−21.4 (−4.1)	−3.9
Discrimination Thermodynamics between Correct and Incorrect Incorporation with No Lesion				
	$-\Delta\Delta H^f$ (kcal/mol)	$-\Delta\Delta G_{37}{}^g$ (kcal/mol)	$-\Delta\Delta S^h$ (eu)	
unmodified -CGA-	−13.8 (−9.1)	−1.6 (−1.7)	−39.3 (−23.9)	

[a] $-\Delta H = -\Delta H_n - (-\Delta H_{n-1})$. [b] $-\Delta G = -\Delta G_n - (-\Delta G_{n-1})$. [c] $-\Delta S = -\Delta S_n - (-\Delta S_{n-1})$. [d] $\Delta T_m = T_{m(n)} - T_{m(n-1)}$. [e] Thermodynamic parameters for TGA are given in parentheses. [f] $-\Delta\Delta H = -\Delta H_{dA\,mismatch} - (-\Delta H_{dC\,match})$. [g] $-\Delta\Delta G_{37} = -\Delta G_{dA\,mismatch} - (-\Delta G_{dC\,match})$. [h] $-\Delta\Delta S = -\Delta S_{dA\,mismatch} - (-\Delta S_{dC\,match})$.

Table 3

Primer Extension with an FAF Lesion at the Primer Terminus (n)				
modified -CG*A-	$-\Delta H^a$ (kcal/mol)	$-\Delta G_{37}{}^b$ (kcal/mol)	$-\Delta S^c$ (eu)	$\Delta T_m{}^d$ (°C)
dC match	−5.9 (0.4)[e]	−0.6 (0.1)	−17.0 (1.2)	−2.9
dA mismatch	3.3 (−1.1)	−0.1 (0.1)	11.0 (−3.7)	−1.6
Discrimination Thermodynamics between Correct and Incorrect Incorporation with an FAF Lesion				
	$-\Delta\Delta H^f$ (kcal/mol)	$-\Delta\Delta G_{37}{}^g$ (kcal/mol)	$-\Delta\Delta S^h$ (eu)	
modified -CG*A-	9.2 (−1.5)	0.5 (0)	28.0 (−4.9)	

[a] $-\Delta H = -\Delta H_n - (-\Delta H_{n-1})$. [b] $-\Delta G = -\Delta G_n - (-\Delta G_{n-1})$. [c] $-\Delta S = -\Delta S_n - (-\Delta S_{n-1})$. [d] $\Delta T_m = T_{m(n)} - T_{m(n-1)}$. [e] Thermodynamic parameters for TG*A are given in parentheses. [f] $-\Delta\Delta H = -\Delta H_{dA\,mismatch} - (-\Delta H_{dC\,match})$. [g] $-\Delta\Delta G_{37} = -\Delta G_{dA\,mismatch} - (-\Delta G_{dC\,match})$. [h] $-\Delta\Delta S = -\Delta S_{dA\,mismatch} - (-\Delta S_{dC\,match})$.

Kf-exo⁻ Bypass Assays. To study the influence of FAF adduction and sequence effect, replication of FAF-modified 16-mer DNA templates (lesion at the first templating position) (Figure 1c) in the presence of all four standard dNTPs and Kf-exo⁻ was studied. Unlike unmodified templates, where primers were readily extended to their full length (data not shown), extension was retarded with modified templates (Figure 4a). A significant amount of the enzyme stalled at the modified template position of CG*A compared to TG*A. In both sequence contexts, insertion opposite the lesion and extension several nucleotides beyond the lesion were also retarded. Doublet bands were observed only at the nearest neighbor downstream base (n + 1) to the lesion in CG*A (arrows). In TG*A, however, additional doublet bands were detected at n + 2 and n + 5 (e.g., −1 deletion 15-mer product).

Single-nucleotide insertion assays using a 20:1 DNA:Kf-exo⁻ ratio were conducted to examine the 5′-sequence effect on the specificity of insertion of a nucleotide opposite FAF. As shown in Figure 4b, in both FAF-modified templates, dCMP was inserted more rapidly opposite FAF than other nucleotides. However, dAMP and dTMP were also incorporated, whereas dGMP was not. At a higher enzyme concentration, however, the extent of insertion of dAMP at the lesion and the two subsequent nucleotide sites was greater in the TG*A than in the CG*A sequence context (Figure 4c).

To examine whether the lesion could influence DNA synthesis prior to the polymerase reaching the lesion, a shorter primer (25-mer) was employed to move the lesion downstream of the synthesis start site. The experiments were performed with the 44-mer template in the presence of all four dNTPs and 2.5 or 5 nM Kf-exo⁻ at 20 °C (Figure 5). While primer extension resulted in full-length products in the presence of unmodified templates (data not shown), significant pausing occurred immediately prior to the lesion (n − 1, 30-mer) and opposite the lesion (n, 31-mer) in both modified sequences. The extent of stalling was greater at the lesion in CG*A (Figure 5a). With TG*A, however, incorporation was blocked more effectively at prelesion site n − 1. In both sequences, it appears that the processivity of Kf-exo⁻ was compromised five to seven bases downstream from the lesion site where the lesion would be positioned in the DNA duplex upstream of the polymerase

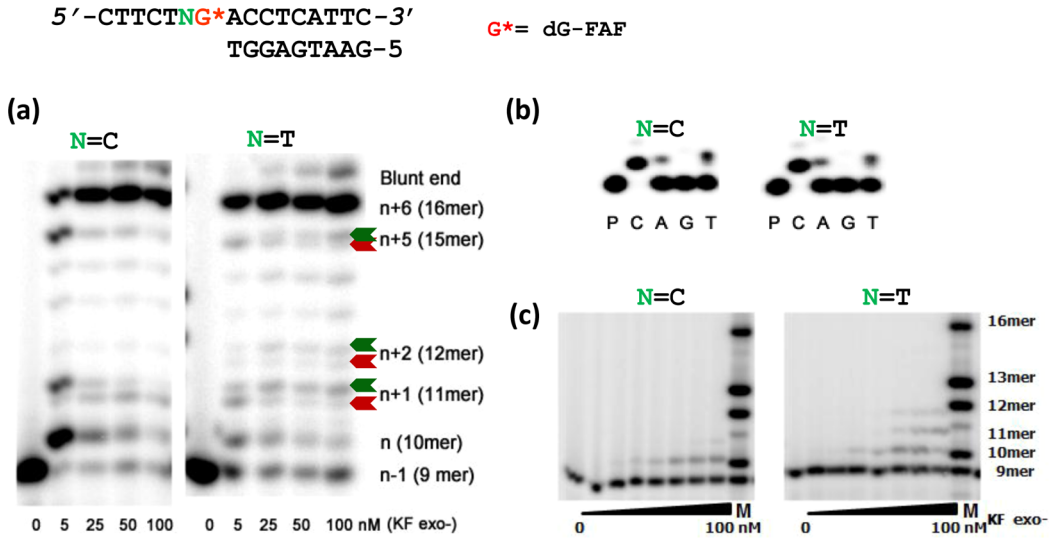


Figure 4. (a) Primer extension assay of a 16/9-mer at different concentrations of Kf-exo⁻. (b) Single-nucleotide incorporation of individual dNTPs opposite FAF-adducted DNA (CG*A and TG*A). (c) Comparison of incorporation of dATP opposite the lesion via variation of the concentration of Kf-exo⁻. The green arrows denote correctly extended products, while red arrows denote dNTP-stabilized misaligned products.

dx.doi.org/10.1021/bi2017443 | Biochemistry 2012, 51, 1983−1995

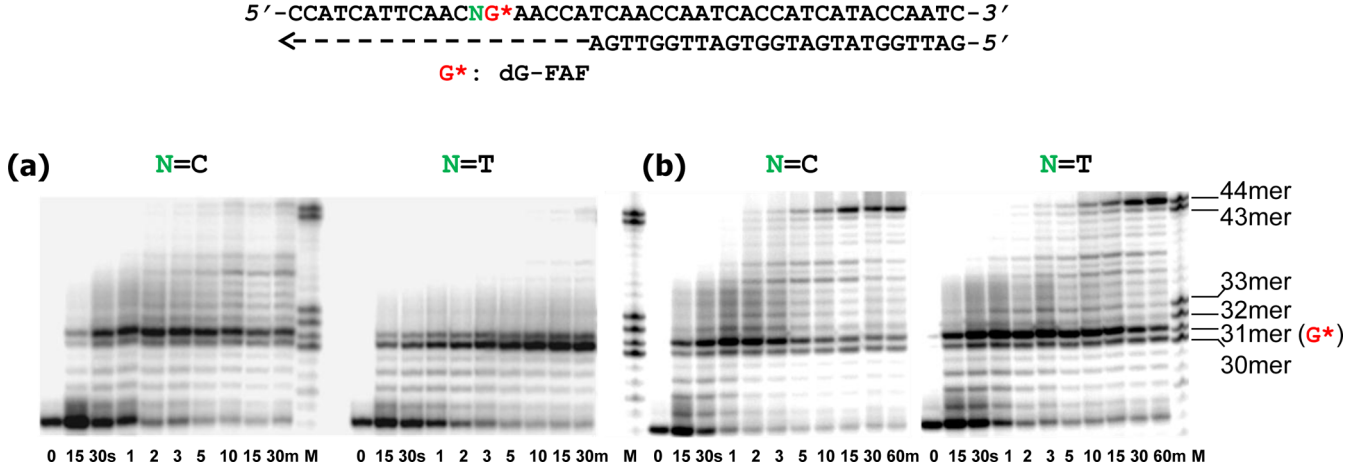


Figure 5. Running start experiments of an aminofluorene-modified DNA (CG*A/TG*A) with all four dNTPs (250 μM each) and (a) 2.5 and (b) 5 nM Kf-exo$^-$ at different time intervals (M, marker).

active site. At a higher enzyme concentration [5.0 nM (Figure 5b)], full-length extension was achieved in both sequences and the differences in pausing behaviors were masked. Consistent with the bypass experiments described above, the extent of polymerase termination was increased in the TG*A sequence in the vicinity of the lesion compared to that for CG*A as seen from the accumulation of products in the vicinity of the lesion (i.e., positions 30 and 31).

Steady-State Kinetics. To gain a better understanding of the influence of FAF-induced conformational heterogeneity on kinetic behavior, steady-state kinetic experiments were performed with the lesion positioned one nucleotide downstream of the templating base (n − 1), at the templating position (n), opposite the primer terminus (n + 1), and three base pairs upstream of the primer terminus (n + 3) (Scheme 1). Additionally, we were interested in whether sequence-specific effects play a role in insertion efficiency (k_{cat}/K_m). Results of the steady-state kinetic experiments are summarized in Tables 4−7.

Table 4. Kinetic Parameters for dTTP Insertion opposite the Templating A One Nucleotide before the Lesion Site (n − 1)

sequence context[a]	k_{cat} (min^{-1})	K_m (μM)	$k_{cat}/K_{m,TTP}$ (μM^{-1} min^{-1})	f_{ins}[b]
-CGA-	11.1 ± 1.7	4.3 ± 3.0	2.6 ± 1.8	1
-CG*A-	3.4 ± 0.2	2.7 ± 0.7	1.3 ± 0.3	0.5
-TGA-	14.7 ± 1.6	1.7 ± 0.9	8.6 ± 4.7	1
-TG*A-	3.4 ± 0.3	1.3 ± 0.5	2.6 ± 1	0.30

[a]The templating base is underlined. [b]$f_{ins} = (k_{cat}/K_{m,dTTP})_{modified}/(k_{cat}/K_{m,dTTP})_{unmodified\ DNA}$.

Insertion Efficiency. With the lesion immediately downstream of the templating nucleotide (n − 1), the relative insertion efficiency (f_{ins}) of the correct nucleotide TTP was hardly affected (f_{ins} = 0.5 and 0.3 with the CG*A and TG*A sequences, respectively) (Table 4) compared to the controls. The f_{ins} of dCMP opposite the lesion in the templating position was reduced considerably (66.1×10^{-4} and 850×10^{-4} for CG*A and TG*A, respectively), which represents 150- and 11-fold rate reductions, respectively, relative to the controls. It is also interesting to note that the insertion frequency of the wrong nucleotide dAMP opposite the adducted guanine in the CG*A context was reduced 10^5-fold compared to 500-fold in the TG*A context, relative to the controls. With CG*A, dCMP insertion was preferred over dAMP ~220-fold, while there was an only 35-fold difference with TG*A (Table 5). The difference in selectivity between the two sequences is primarily due to the lower $K_{m,dATP}$ observed with the TG*A sequence, suggesting that the downstream templating thymine is influencing selectivity. In other words, insertion of dATP may use the downstream thymidine as the templating nucleotide via a dNTP-stabilized misalignment. The apparent $K_{m,dATP}$ in this sequence context (5′-TG) is hardly affected by the lesion (decreased 5−6-fold relative to that with dCTP) (Table 5). In contrast, the apparent $K_{m,dATP}$ in the 5′-CG sequence context is increased (40−70-fold), suggesting that in this situation the downstream templating base (i.e., C) does not serve as a coding nucleotide.

Extension Efficiency. At n + 1, the extension efficiency (correct insertion efficiency beyond the lesion site) between CG*A and TG*A differed by 58- and 4-fold, respectively, in the dC-match series (Table 6). With dA at the primer terminus positioned opposite G*, correct insertion at site n + 1 resulted in 10^5- and 10^4-fold reductions for CG*A and TG*A, respectively. However, the impact of the lesion on the extension frequency at n + 3 was minimal as the extension frequency was 88×10^{-2} and 48×10^{-2} for CG*A and TG*A, respectively, in the dC-match series (Table 7). In the dA-mismatch series, the extension efficiency differed by only 17-fold between the CG*A and TG*A sequences.

DISCUSSION

It is well-known that DNA sequence is a major determining factor for mutational and repair outcomes of site-specifically modified bulky DNA lesions.[4,31] In this study, we investigated the impact of lesion-induced conformational heterogeneity on primer extension during a simulated TLS. Specifically, we probed the 5′-base sequence effect (5′-TG*A vs 5′-CG*A) on aminofluorene-induced S/B/W heterogeneity and nucleotide insertion efficiencies catalyzed by Kf-exo$^-$.

Conformational Effect of 5′-Flanking Sequence and Mismatch on the FAF-Modified TLS System. The ^{19}F NMR/ICD characteristics of the simulated TLS systems involving the FAF-modified 16-mer TG*A[29] and CG*A templates were similar. At prelesion site n − 1, the lesion adopted a dynamic mixture of *syn*-S* and *anti*-B* conformers

dx.doi.org/10.1021/bi2017443 | *Biochemistry* 2012, 51, 1983−1995

Table 5. Kinetic Parameters for Insertion opposite the Lesion (n)

sequence context[a]	incoming dNTP	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1} min^{-1})	f_{ins}
-CGA-	dCTP	7.31 ± 0.9	0.2 ± 0.1	36.6 ± 18.8	1
	dATP	0.15 ± 0.02	13.4 ± 5.0	0.011 ± 0.004	3.00×10^{-4}
-CG*A-	dCTP	0.8 ± 0.16	3.3 ± 3.2	0.242 ± 0.240	66.1×10^{-4}
	dATP	0.14 ± 0.03	126 ± 55	0.0011 ± 0.0005	0.30×10^{-4}
-TGA-	dCTP	11.66 ± 1.5	4.7 ± 2.8	2.5 ± 1.5	1
	dATP	0.14 ± 0.01	28.3 ± 9.8	0.0049 ± 0.0017	19.6×10^{-4}
-TG*A-	dCTP	0.34 ± 0.03	1.6 ± 0.7	0.2125 ± 0.0948	850×10^{-4}
	dATP	0.05 ± 0.004	8.3 ± 3.8	0.0060 ± 0.00028	24.0×10^{-4}

[a]The templating base is underlined.

Table 6. Kinetic Parameters for Correct (dGTP or dATP) Insertion with the Template Lesion opposite the Primer Terminus ($n + 1$)

sequence context[a]	template–primer terminus	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1} min^{-1})	f_{ext}[b]
-CGA-	G:C	19.7 ± 0.9	2.4 ± 0.5	8.2 ± 1.7	1
	G:A	0.018 ± 0.002	12.7 ± 5.1	0.0014 ± 0.0006	1.71×10^{-4}
-CG*A-	G*:C	1.24 ± 0.17	8.9 ± 3.3	0.139 ± 0.055	170×10^{-4}
	G*:A	0.004 ± 0.001	15.6 ± 10.7	0.00026 ± 0.00019	0.32×10^{-4}
-TGA-	G:C	18.1 ± 3.6	17.5 ± 10.5	1.0 ± 0.7	1
	G:A	0.022 ± 0.002	28 ± 5.6	0.00079 ± 0.00017	7.9×10^{-4}
-TG*A-	G*:C	0.77 ± 0.10	2.9 ± 1.9	0.266 ± 0.177	2660×10^{-4}
	G*:A	0.006 ± 0.002	46 ± 31	0.00013 ± 0.0001	1.30×10^{-4}

[a]The templating base is underlined. [b]$f_{ext} = (k_{cat}/K_m)_{mismatch\ or\ modified\ terminus}/(k_{cat}/K_m)_{unmodified}$.

Table 7. Kinetic Parameters for Correct TTP Insertion with the Template Lesion 3 bp Upstream of the Primer Terminus ($n + 3$)

sequence context	nucleotide opposite G or G*	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1} min^{-1})	f_{ext}[a]
-CGA-	C	32 ± 13	75 ± 56	0.43 ± 0.36	1
	A	1.0 ± 0.4	46 ± 32	0.021 ± 0.017	4.9×10^{-2}
-CG*A-	C	5.0 ± 0.5	13.2 ± 4.4	0.379 ± 0.132	88×10^{-2}
	A	0.47 ± 0.08	20.6 ± 10.8	0.023 ± 0.013	5.3×10^{-2}
-TGA-	C	18.0 ± 3.6	77.9 ± 29.3	0.231 ± 0.098	1
	A	0.020 ± 0.001	53 ± 5	0.00038 ± 0.00004	0.16×10^{-2}
-TG*A-	C	0.21 ± 0.05	1.9 ± 0.3	0.111 ± 0.032	48×10^{-2}
	A	0.008 ± 0.002	12 ± 10	0.0007 ± 0.0006	0.3×10^{-2}

[a]$f_{ext} = (k_{cat}/K_m)_{mismatch\ or\ modified\ terminus}/(k_{cat}/K_m)_{unmodified}$.

(see below). Conformational profiles for the remaining TLS extension (e.g., n to $n + 6$) were influenced by the nature of the inserted nucleotide at lesion site n and the extent of primer extension. Incorporation of the correct base, dC, at lesion site n and subsequent replication resulted in a gradual transition from the flexible *syn*-S* to an S/B conformer mixture. A transition occurred in the minor groove binding the *syn*-W* conformer for the dA-mismatch series. Nevertheless, there were subtle spectral differences between the two sequence contexts. For example, the signal of the B conformer of the $n + 6$ CG*A duplex was split and coalesced at low temperatures [15–25 °C (Figure 2)], whereas the corresponding TG*A duplex showed no such splitting.[29]

Perhaps the most notable result was a difference in the S:B population ratios for fully paired $n + 6$ duplexes in the dC-match series, i.e., 58:42 and 38:62 for the CG*A and TG*A sequences, respectively. The results were in close agreement with those observed for FAF-modified 12-mer duplexes with the same flanking sequence contexts reported previously: 56:44 and 45:55, respectively.[16] Of particular interest, with regard to this work, are the AF- and 2-amino-3-methylimidazo[4,5-*f*]-quinolone (IQ)-modified duplex adducts at each of the three dG bases (5′-C$\underline{G_1}\underline{G_2}C\underline{G_3}$CC-3′) in the *Nar*I sequence, a well-known mutational hot spot for arylamine mutagens in *E. coli*.[32−34] The S:B ratios for AF at G_1, G_2, and G_3 have been determined to be 30:70, 10:90, and 50:50, respectively.[35] Similar studies with IQ showed that G_1 and G_2 duplexes adopt a minor groove-binding W-type conformer, whereas the G_3 duplex assumes exclusively the S conformation. The authors have suggested that the hydrophobic quinoline ring might stack favorably with the flanking dG rather than dC in the complementary strand.[34] A similar case could be made for the aforementioned highly S conformer of the G_3-AF duplex. Similarly, a fully paired 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP)-modified 11-mer duplex with the CG*C context adopts largely the S-type conformation.[36] These results, taken together, indicate that C on the 5′-side of an arylamine lesion promotes the S conformation over the B conformation, consistent with our findings.

As mentioned, lesion flexibility would not allow definition of a clear S/B heterogeneity at the lesion (n) and immediate neighboring ($n - 1$ and $n + 1$) sites. In both sequence contexts, the exposed nature of the aminofluorene moiety is evident from broad ^{19}F NMR signals, large H/D isotope effects,[37] and a large blob of signals observed in the imino proton spectra.[29] These results, coupled with a strongly positive $ICD_{290-350}$, suggested a

dx.doi.org/10.1021/bi2017443 | *Biochemistry* 2012, 51, 1983−1995

mixture of *syn*-S* and *anti*-B* conformers with predominantly *syn* conformers. Previous ^{1}H NMR studies have shown that AF at the replication fork (n − 1) shares the features of the S conformation, regardless of whether there is a dC partner, a mismatched dA, or no base opposite the lesion.[38,39] The FAF-modified dG is displaced into the major groove, while the carcinogen is inserted in its place. Subtle differences exist in the rotameric possibilities about the β'-linkage and the complementary base at the lesion site. However, these data were obtained in buffer-only solutions, thus lacking polymerase-induced conformational consideration.[11,40] Dutta et al.[17] have shown fuzzy electron densities around the lesion in the active site AF-bound T7 DNA polymerase, which implies a lesion-induced conformational heterogeneity in the solid state as well. Nonetheless, we theorize that heterogeneity at the lesion site can be extrapolated from the well-defined S/B heterogeneity of the fully paired n + 6 duplexes. As such, it is assumed that a greater S conformation of CG*A over TG*A exists.

Correlation of Kf-exo$^-$ Activity with S/B Conformational Heterogeneity and Thermodynamics. There was a stronger block to Kf-exo$^-$-dependent DNA synthesis at the site of the lesion (n) in CG*A than in TG*A (Figure 4a). Incorporation of dC into the CG*A system resulted in substantial reductions in $-\Delta H$ (−5.9 kcal/mol) and $-\Delta G_{37\,°C}$ (−0.6 kcal/mol). This is in contrast to the values reported previously for the TG*A series, which exhibited small increases in $-\Delta H$ (0.4 kcal/mol) and $-\Delta G_{37\,°C}$ (0.1 kcal/mol).[28] The loss of thermodynamic stability of CG*A relative to TG*A may be related to the greater population of the S conformer in the CG*A system, which hinders Watson−Crick base pairing at the lesion site.

Unlike the controls that produced full primer extension, significant pausing occurred at the n − 1 (30-mer) and n (31-mer) sites for modified templates. The extent of stalling was greater at the lesion in CG*A, whereas incorporation was blocked more effectively and persistently at prelesion site n − 1 of TG*A (Figure 5a). As mentioned above, the block at the lesion site could be related to the greater presence of the S conformer. Kinetic results in this study (Table 4) and previous studies[25] showed that Kf-exo$^-$ is not highly discriminatory with respect to the prelesion areas, including the n − 1 site. The persistent blocking before the lesion site in TG*A could be caused by weaker discrimination between the correct and wrong nucleotide insertion opposite the lesion site as the insertion frequency differed by only 35-fold.

Crystallographic studies by Hsu et al.[41] presented snapshot structures of how the AF adduct could be accommodated in the active site of the *Bacillus* fragment and undergoes one round of replication. At lesion site n, the modified dG in the preinsertion pocket adopts the *syn* conformation, with the carcinogen ring experiencing extensive van der Waals interactions with the nearby O helices. In transiting to the postinsertion pocket, the modified dG rotates to the *anti* conformation with the carcinogen in the major groove pocket (B conformer) preparing to form stable Watson−Crick base pairs with an incoming dCTP. This accounts for the single-nucleotide insertion assay data, which showed predominant insertion of dCTP opposite FAF (Figure 4b). The insertion assays also showed a small but significant insertion of dATP opposite FAF (Figure 4c), and the extent appeared to be sensitive to the nature of the base 5′ to the lesion. Previous site-specific mutagenesis experiments in simian cells have revealed that incorporation of dATP ranged from 2 to 4% with cytosine, adenine, and thymine template bases placed 5′ to the lesion to 70% with a guanine base placed 5′ to the lesion.[5] Misinsertion of dATP at the lesion site would result in G → T transversion mutations, which are the most frequently observed mutations for AF in vivo. In principle, the B conformer of FAF in the active site can accommodate any dNTPs; however, only base pairs with dCTP would attain a catalytically competent closed conformation. The S conformer of FAF could clash with the incoming dATP to favor a W-type conformation, in which the aminofluorene moiety is wedged and stabilized in the minor groove region. This scenario is plausible because of a decrease in entropy in the dA-mismatch series compared to the dC-match series where a gain in entropy was observed. Previous NMR studies have suggested the existence of such an AF-G:A mismatch by way of one hydrogen bond between the protonated N1 atom of dA and O^6 of modified dG.[42] However, because polymerases tend to interact with the DNA minor groove, primer extension could be terminated with the W conformation.

According to the DSC-based thermodynamic data (Table 1), addition of dC in the control CGA template sequence was favored over insertion of dA ($-\Delta\Delta H$ = −13.8 kcal/mol, and $-\Delta\Delta G°$ = −1.6 kcal/mol) with considerable enthalpy−entropy compensation (Table 2). Similar values were obtained for the TGA control template ($-\Delta\Delta H$ = −9.1 kcal/mol, and $-\Delta\Delta G°$ = −1.7 kcal/mol).[28] Formation of regular Watson−Crick base pairs results in a typical additive pattern of thermodynamics with primer extension. With the FAF-modified CG*A sequence, however, insertion of dA over dC was preferred enthalpically ($-\Delta H$ = −5.9 kcal/mol over 3.3 kcal/mol, and $-\Delta\Delta H$ = 9.2 kcal/mol). This is in contrast to the TG*A series, which resulted in comparable enthalpies regardless of whether dC or dA was inserted at the lesion site [$-\Delta\Delta H$ = −1.5 kcal/mol (Table 3)]. The unusual positive thermodynamic stability ($-\Delta\Delta H$ = 9.2 kcal/mol) of dA over dC of CG*A is not consistent with the kinetic results that showed a ∼220-fold preferential insertion of dCTP over dATP. It should also be noted that the insertion efficiency of dAMP opposite FAF was found to be 10-fold greater in TG*A than in CG*A (Figure 6).

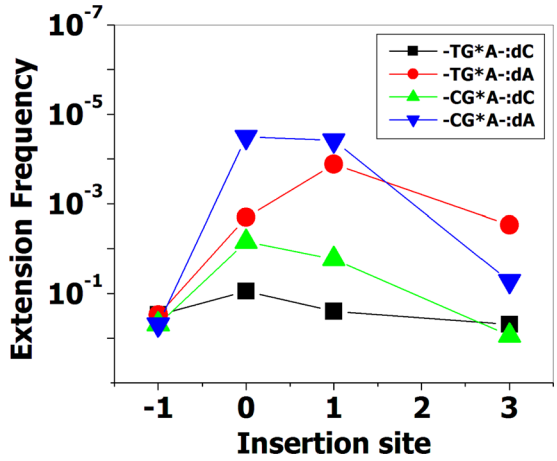


Figure 6. Plots of extension frequency vs insertion site.

It has been shown that 5′-T consistently promotes incorporation of dATP opposite 8-oxo-G, 3-(2′-deoxy-β-D-*erythro*-pentofuranosyl)-pyrimido[1,2-*a*]purin-10(3*H*)-one (M1dG), an abasic site, and Fapy adducts.[43−46] In earlier studies, insertion of the wrong nucleotide was proposed either by direct misinsertion opposite the lesion or by misaligned primer−template strands in which the lesion was looped out of the helix

dx.doi.org/10.1021/bi2017443 | *Biochemistry* 2012, 51, 1983−1995

and favored the nucleotide insertion complementary to the 5′-base positioned to the lesion. In this study, similar pathways cannot be ruled out as $K_{m,dATP}$ values were similar at sites n and $n + 1$. In addition, doublet bands were observed in the bypass experiments (see Figure 4a in Results). A similar misalignment mechanism with Kf and an abasic site has been reported in which the incoming nucleotide complementary to a 5′-base downstream to the lesion site was favored.[45]

Long-Range Effect of the AF-Induced S/B Conformational Heterogeneity in Translesion DNA Synthesis. Primer extension in the $n - 1$ to $n + 2$ extension of the dC-match series negatively impacted the thermodynamic parameters, but a usual additive gain pattern of T_m and enthalpy was resumed at $n + 3$ (Figure 3). This long-range effect is consistent with the bypass data shown in Figure 5, which depicts compromised Kf-exo$^-$ activities several bases downstream of the lesion site. The results are also in accordance with the primer extension kinetics (Tables 6 and 7). The efficiency of correct nucleotide insertion was significantly reduced when the lesion was at the primer terminus or 3 bp upstream of the primer terminus. The impact was generally 100-fold greater for CG*A than for TG*A (Figure 6). As postulated above, this may also be due to the greater population of the S conformer of the CG*A over the TG*A series. Within each sequence, however, the relative insertion efficiencies for the dA-mismatch series were consistently slower compared to those in the dC-match series. It is likely that the flexible *syn*-S* conformer FAF at lesion site n in the dA-mismatch series (see above) may interfere with the ability of the incoming dAMP to form Watson−Crick base pairs, thereby reducing the overall replication fidelity. While there was considerable entropy compensation (Table 1), the data showed that the enthalpy terms are primarily responsible for the adduct-induced thermal and thermodynamic destabilization.[47] The greater impact on mutagenic frequency for the dA-mismatch series versus the dC-match series is in good agreement with the steady-state primer kinetic results reported by Miller and Grollman[25] on various mutagenic adducts, including AF and AAF. Johnson and Beese[48] have shown that mismatch-induced distortion is not just localized to the mismatch site but extends several base pairs downstream through the so-called "DNA binding region" of the *Bacillus* fragment (Bf) DNA polymerase, a high-fidelity DNA polymerase that is extensively homologous in sequence and structure to Kf.[19]

We have previously modeled the rounds of replication for the B and W conformers of AF adducts at various primer positions (n to $n + 3$) docked in the active site of the Bf.[19] The hydrophobic aminofluorene moiety in the minor groove of the W conformer would impose a major steric clash with the DNA binding area of the polymerase, thus resulting in a dramatic reduction in the replication rate. This is reminiscent of the minor groove binding (+)-*trans*-*anti*-BP−DNA adduct, in which the bulky benzo[*a*]pyrene imposes a major disruption in the duplex binding site.[49,50] A similar effect has been observed for (*R*)- and (*S*)-N^2-styrene oxide minor groove adducts, which pose an impediment for HIV-1 reverse transcriptase.[51] As for the B-type conformer, the carcinogenic moiety orients into the energetically favorable solvent-exposed major groove, thus causing less disruption of replication as the DNA binding region translocates through the enzyme. However, the B conformer may perturb the dynamic and geometric constraints of the grooves, which may slow replication rates for the downstream bases to the lesion as the bound DNA substrate has to undergo the B- to A-form transition approximately four nucleotides from the template−primer junction. At the further downstream $n + 6$ position, however, the AF lesion is located near the tip of the duplex binding region of the polymerase binding region and therefore exerts virtually no adverse protein−DNA interactions. The dynamic NMR results and primer extension kinetic data presented here support a model in which adduct-induced conformational heterogeneities at positions remote from the replication fork affect polymerase function through a long-range protein−DNA interaction.

In summary, our data revealed that (a) AF in the CG*A duplex sequence adopts a greater population of S conformer than TG*A sequence, (b) the S conformer of CG*A thermodynamically favors insertion of A over C at the lesion site (n), (c) significant stalling occurred at both the prelesion ($n - 1$) and lesion (n) sites but the effect was more persistent for the S conformer of CG*A than TG*A at the lesion site (n), and (d) relative nucleotide insertion frequencies (f_{ins}) were greater for the B conformer of TG*A for either dCTP or dATP at the lesion site (n) and the insertion rate was significantly reduced at n and immediate downstream base pairs ($n + 1$ and $n + 3$). Taken together, these results emphasize the importance of lesion-induced conformational heterogeneity in translesion DNA synthesis. To the best of our knowledge, this work represents a novel structure−function relationship in which adduct structure is directly linked to nucleotide insertion efficiency in a conformation-specific manner during translesion synthesis. Importantly, this study reinforces Loechler's original hypothesis,[4] which states that DNA adducts may exist in multiple conformations in the active site of a polymerase, each contributing to a different mutational outcome.

■ ASSOCIATED CONTENT

Supporting Information

HPLC and UV profile of the FAF-modified oligonucleotide, ^{1}H imino proton NMR and CD spectral series of simulated translesion synthesis, and plots of ΔH versus insertion site, and ΔG versus insertion site. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: bcho@uri.edu. Telephone: (401) 874-5024. Fax: (401) 874-5766.

Funding

We are grateful to the National Institutes of Health (R01CA098296) for their financial support of this work. This research was also made possible in part by the use of the Research Core Facility supported by the National Center for Research Resources (P20 RR016457).

■ ACKNOWLEDGMENTS

We thank Dr. William Beard for his valuable input and critical reading of the manuscript.

■ ABBREVIATIONS

AF, *N*-(2′-deoxyguanosin-8-yl)-2-aminofluorene; FAF, *N*-(2′-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene; DSC, differential scanning calorimetry; ^{19}F NMR, fluorine nuclear magnetic resonance spectroscopy; $ICD_{290-350}$, induced circular dichroism at 290−350 nm; TLS, translesion DNA synthesis.

REFERENCES

(1) Luch, A. (2005) Nature and nurture: Lessons from chemical carcinogenesis. *Nat. Rev. Cancer 5*, 113−125.

(2) Hubscher, U., and Maga, G. (2011) DNA replication and repair bypass machines. *Curr. Opin. Chem. Biol. 5*, 627−635.

(3) Kelly, M. R. (2011) *DNA Repair in Cancer Therapy*, Academic Press, Amsterdam.

(4) Seo, K.-Y., Jelinsky, S. A., and Loechler, E. L. (2000) Factors that influence the mutagenic patterns of DNA adducts from chemical carcinogens. *Mutat. Res. 463*, 215−246.

(5) Shibutani, S., Suzuki, N., Tan, X., Johnson, F., and Grollman, A. P. (2001) Influence of Flanking Sequence Context on the Mutagenicity of Acetylaminofluorene-Derived DNA Adducts in Mammalian Cells. *Biochemistry 40*, 3717−3722.

(6) Skipper, P. L., Kim, M. Y., Sun, H. L., Wogan, G. N., and Tannenbaum, S. R. (2010) Monocyclic aromatic amines as potential human carcinogens: Old is new again. *Carcinogenesis 31*, 50−58.

(7) Neumann, H. G. (2007) Aromatic amines in experimental cancer research: Tissue-specific effects, an old problem and new solutions. *Crit. Rev. Toxicol. 37*, 211−236.

(8) Turesky, R. J., and Le Marchand, L. (2011) Metabolism and biomarkers of heterocyclic aromatic amines in molecular epidemiology studies: Lessons learned from aromatic amines. *Chem. Res. Toxicol. 24*, 1169−1214.

(9) Heflich, R. H., and Neft, R. E. (1994) Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. *Mutat. Res. 318*, 73−174.

(10) Beland, F. A., and Kadlubar, F. F. (1990) *Handbook of Experimental Pharmacology*, pp 267−325, Spring-Verlag, Heidelberg, Germany.

(11) Patel, D. J., Mao, B., Gu, Z., Hingerty, B. E., Gorin, A., Basu, A. K., and Broyde, S. (1998) Nuclear magnetic resonance solution structures of covalent aromatic amine-DNA adducts and their mutagenic relevance. *Chem. Res. Toxicol. 11*, 391−407.

(12) Cho, B. (2010) Structure-function characteristics of aromatic amine-DNA adducts. In *The Chemical Biology of DNA Damage*, pp 217−238, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.

(13) Lukin, M., and de Los Santos, C. (2006) NMR structures of damaged DNA. *Chem. Rev. 106*, 607−686.

(14) Cho, B. P. (2004) Dynamic conformational heterogeneities of carcinogen-DNA adducts and their mutagenic relevance. *J. Environ. Sci. Health, Part C: Environ. Carcinog. Ecotoxicol. Rev. 22*, 57−90.

(15) Patnaik, S., and Cho, B. P. (2010) Structures of 2-acetylaminofluorene modified DNA revisited: Insight into conformational heterogeneity. *Chem. Res. Toxicol. 23*, 1650−1652.

(16) Meneni, S. R., Shell, S. M., Gao, L., Jurecka, P., Lee, W., Sponer, J., Zou, Y., Chiarelli, M. P., and Cho, B. P. (2007) Spectroscopic and theoretical insights into sequence effects of aminofluorene-induced conformational heterogeneity and nucleotide excision repair. *Biochemistry 46*, 11263−11278.

(17) Dutta, S., Li, Y., Johnson, D., Dzantiev, L., Richardson, C. C., Romano, L. J., and Ellenberger, T. (2004) Crystal structures of 2-acetylaminofluorene and 2-aminofluorene in complex with T7 DNA polymerase reveal mechanisms of mutagenesis. *Proc. Natl. Acad. Sci. U.S.A. 101*, 16186−16191.

(18) Jain, N., Meneni, S., Jain, V., and Cho, B. P. (2009) Influence of flanking sequence context on the conformational flexibility of aminofluorene-modified dG adduct in dA mismatch DNA duplexes. *Nucleic Acids Res. 37*, 1628−1637.

(19) Meneni, S., Liang, F., and Cho, B. P. (2007) Examination of the Long-range Effects of Aminofluorene-induced Conformational Heterogeneity and Its Relevance to the Mechanism of Translesional DNA Synthesis. *J. Mol. Biol. 366*, 1387−1400 , Erratum, volume 1398, p 1786.

(20) Nagalingam, A., Seo, K. Y., and Loechler, E. L. (2005) Mutagenesis studies of the major benzo[a]pyrene N2-dG adduct in a 5′-TG versus a 5′-UG sequence: Removal of the methyl group causes a modest decrease in the [G->T/G->A] mutational ratio. *Mutagenesis 20*, 105−110.

(21) Meneni, S., Shell, S. M., Zou, Y., and Cho, B. P. (2007) Conformation-specific recognition of carcinogen-DNA adduct in *Escherichia coli* nucleotide excision repair. *Chem. Res. Toxicol. 20*, 6−10.

(22) Zou, Y., Shell, S. M., Utzat, C. D., Luo, C., Yang, Z., Geacintov, N. E., and Basu, A. K. (2003) Effects of DNA adduct structure and sequence context on strand opening of repair intermediates and incision by UvrABC nuclease. *Biochemistry 42*, 12654−12661.

(23) Ruan, Q., Liu, T., Kolbanovskiy, A., Liu, Y., Ren, J., Skorvaga, M., Zou, Y., Lader, J., Malkani, B., Amin, S., Van Houten, B., and Geacintov, N. E. (2007) Sequence context- and temperature-dependent nucleotide excision repair of a benzo[a]pyrene diol epoxide-guanine DNA adduct catalyzed by thermophilic UvrABC proteins. *Biochemistry 46*, 7006−7015.

(24) Cai, Y., Patel, D. J., Geacintov, N. E., and Broyde, S. (2007) Dynamics of a Benzo[a]pyrene-derived Guanine DNA Lesion in TGT and CGC Sequence Contexts: Enhanced Mobility in TGT Explains Conformational Heterogeneity, Flexible Bending, and Greater Susceptibility to Nucleotide Excision Repair. *J. Mol. Biol. 374*, 292−305.

(25) Miller, H., and Grollman, A. P. (1997) Kinetics of DNA polymerase I (Klenow fragment exo$^-$) activity on damaged DNA templates: Effect of proximal and distal template damage on DNA synthesis. *Biochemistry 36*, 15336−15342.

(26) Rechkoblit, O., Kolbanovskiy, A., Malinina, L., Geacintov, N. E., Broyde, S., and Patel, D. J. (2010) Mechanism of error-free and semitargeted mutagenic bypass of an aromatic amine lesion by Y-family polymerase Dpo4. *Nat. Struct. Mol. Biol. 17*, 379−388.

(27) Suzuki, N., Ohashi, E., Hayashi, K., Ohmori, H., Grollman, A. P., and Shibutani, S. (2001) Translesional synthesis past acetylaminofluorene-derived DNA adducts catalyzed by human DNA polymerase κ and *Escherichia coli* DNA polymerase IV. *Biochemistry 40*, 15176−15183.

(28) Liang, F., and Cho, B. P. (2007) Probing the thermodynamics of aminofluorene-induced translesion DNA synthesis by differential scanning calorimetry. *J. Am. Chem. Soc. 129*, 12108−12109.

(29) Liang, F., and Cho, B. P. (2011) Conformational and thermodynamic impact of bulky aminofluorene adduction on simulated translesion DNA synthesis. *Chem. Res. Toxicol. 24*, 597−605.

(30) Liang, F., Meneni, S., and Cho, B. P. (2006) Induced circular dichroism characteristics as conformational probes for carcinogenic aminofluorene-DNA adducts. *Chem. Res. Toxicol. 19*, 1040−1043.

(31) Cai, Y., Patel, D. J., Broyde, S., and Geacintov, N. E. (2010) Base sequence context effects on nucleotide excision repair. *J. Nucleic Acids 2010*, 174252.

(32) Elmquist, C. E., Wang, F., Stover, J. S., Stone, M. P., and Rizzo, C. J. (2007) Conformational differences of the C8-deoxyguanosine adduct of 2-amino-3-methylimidazo[4,5-f]quinoline (IQ) within the NarI recognition sequence. *Chem. Res. Toxicol. 20*, 445−454.

(33) Wang, F., DeMuro, N. E., Elmquist, C. E., Stover, J. S., Rizzo, C. J., and Stone, M. P. (2006) Base-displaced intercalated structure of the food mutagen 2-amino-3-methylimidazo[4,5-f]quinoline in the recognition sequence of the NarI restriction enzyme, a hotspot for −2 bp deletions. *J. Am. Chem. Soc. 128*, 10085−10095.

(34) Wang, F., Elmquist, C. E., Stover, J. S., Rizzo, C. J., and Stone, M. P. (2007) DNA sequence modulates the conformation of the food mutagen 2-amino-3-methylimidazo[4,5-f]quinoline in the recognition sequence of the NarI restriction enzyme. *Biochemistry 46*, 8498−8516.

(35) Jain, N., Li, Y., Zhang, L., Meneni, S. R., and Cho, B. P. (2007) Probing the sequence effects on NarI-induced −2 frameshift mutagenesis by dynamic ^{19}F NMR, UV, and CD spectroscopy. *Biochemistry 46*, 13310−13321.

(36) Brown, K., Hingerty, B. E., Guenther, E. A., Krishnan, V. V., Broyde, S., Turteltaub, K. W., and Cosman, M. (2001) Solution structure of the 2-amino-1- methyl-6-phenylimidazo[4,5-b]pyridine C8-deoxyguanosine adduct in duplex DNA. *Proc. Natl. Acad. Sci. U.S.A. 98*, 8507−8512.

(37) Zhou, L., Rajabzadeh, M., Traficante, D. D., and Cho, B. P. (1997) Conformational Heterogeneity of Arylamine-Modified DNA: ^{19}F NMR Evidence. *J. Am. Chem. Soc. 119*, 5384−5389.
(38) Gu, Z., Gorin, A., Hingerty, B. E., Broyde, S., and Patel, D. J. (1999) Solution structures of aminofluorene [AF]-stacked conformers of the syn [AF]-C8-dG adduct positioned opposite dC or dA at a template-primer junction. *Biochemistry 38*, 10855−10870.
(39) Mao, B., Gu, Z., Gorin, A., Hingerty, B. E., Broyde, S., and Patel, D. J. (1997) Solution structure of the aminofluorene-stacked conformer of the syn [AF]-C8-dG adduct positioned at a template-primer junction. *Biochemistry 36*, 14491−14501.
(40) Broyde, S., Wang, L., Zhang, L., Rechkoblit, O., Geacintov, N. E., and Patel, D. J. (2008) DNA adduct structure-function relationships: Comparing solution with polymerase structures. *Chem. Res. Toxicol. 21*, 45−52.
(41) Hsu, G. W., Kiefer, J. R., Burnouf, D., Becherel, O. J., Fuchs, R. P. P., and Beese, L. S. (2004) Observing Translesion Synthesis of an Aromatic Amine DNA Adduct by a High-fidelity DNA Polymerase. *J. Biol. Chem. 279*, 50280−50285.
(42) Norman, D., Abuaf, P., Hingerty, B. E., Live, D., Grunberger, D., Broyde, S., and Patel, D. J. (1989) NMR and computational characterization of the N-(deoxyguanosin-8-yl)aminofluorene adduct [(AF)G] opposite adenosine in DNA: (AF)G[syn]·A[anti] pair formation and its pH dependence. *Biochemistry 28*, 7462−7476.
(43) Jaloszynski, P., Ohashi, E., Ohmori, H., and Nishimura, S. (2005) Error-prone and inefficient replication across 8-hydroxyguanine (8-oxoguanine) in human and mouse ras gene fragments by DNA polymerase κ. *Genes Cells 10*, 543−550.
(44) Maddukuri, L., Eoff, R. L., Choi, J. Y., Rizzo, C. J., Guengerich, F. P., and Marnett, L. J. (2010) In vitro bypass of the major malondialdehyde- and base propenal-derived DNA adduct by human Y-family DNA polymerases κ, ι, and Rev1. *Biochemistry 49*, 8415−8424.
(45) Berthet, N., Roupioz, Y., Constant, J. F., Kotera, M., and Lhomme, J. (2001) Translesional synthesis on DNA templates containing the 2′-deoxyribonolactone lesion. *Nucleic Acids Res. 29*, 2725−2732.
(46) Asagoshi, K., Terato, H., Ohyama, Y., and Ide, H. (2002) Effects of a guanine-derived formamidopyrimidine lesion on DNA replication: translesion DNA synthesis, nucleotide insertion, and extension kinetics. *J. Biol. Chem. 277*, 14589−14597.
(47) Liang, F., and Cho, B. P. (2010) Enthalpy-entropy contribution to carcinogen-induced DNA conformational heterogeneity. *Biochemistry 49*, 259−266.
(48) Johnson, S. J., and Beese, L. S. (2004) Structures of mismatch replication errors observed in a DNA polymerase. *Cell 116*, 803−816.
(49) Hsu, G. W., Huang, X., Luneva, N. P., Geacintov, N. E., and Beese, L. S. (2005) Structure of a high fidelity DNA polymerase bound to a benzo[a]pyrene adduct that blocks replication. *J. Biol. Chem. 280*, 3764−3770.
(50) Alekseyev, Y. O., and Romano, L. J. (2002) Effects of benzo[a]pyrene adduct stereochemistry on downstream DNA replication in vitro: Evidence for different adduct conformations within the active site of DNA polymerase I (Klenow fragment). *Biochemistry 41*, 4467−4479.
(51) Forgacs, E., Latham, G., Beard, W. A., Prasad, R., Bebenek, K., Kunkel, T. A., Wilson, S. H., and Lloyd, R. S. (1997) Probing structure/function relationships of HIV-1 reverse transcriptase with styrene oxide N2-guanine adducts. *J. Biol. Chem. 272*, 8525−8530.