BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

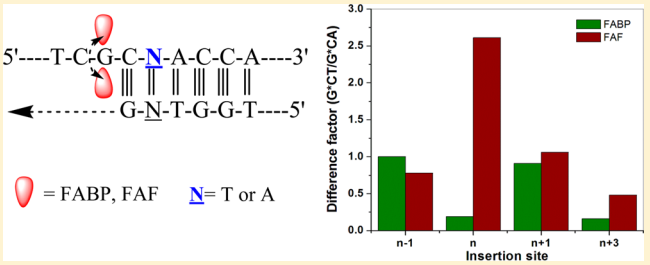
Conformational Insights into the Lesion and Sequence Effects for Arylamine-Induced Translesion DNA Synthesis: ^{19}F NMR, Surface Plasmon Resonance, and Primer Kinetic Studies

Vipin Jain, Vaidyanathan G. Vaidyanathan,[†] Satyakam Patnaik,[‡] Sathyaraj Gopal, 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: Adduct-induced DNA damage can affect transcription efficiency and DNA replication and repair. We previously investigated the effects of the 3′-next flanking base (G*C*T* vs G*C*A*; G*, FABP, *N*-(2′-deoxyguanosin-8-yl)-4′-fluoro-4-aminobiphenyl; FAF, *N*-(2′-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene) on the conformation of arylamine-DNA lesions in relation to *E. coli* nucleotide excision repair (Jain, V., Hilton, B., Lin, B., Patnaik, S., Liang, F., Darian, E., Zou, Y., Mackerell, A. D., Jr., and Cho, B. P. (2013) *Nucleic Acids Res., 41*, 869−880). Here, we report the differential effects of the same pair of sequences on DNA replication *in vitro* by the polymerases exofree Klenow fragment (Kf-exo$^-$) and Dpo4. We obtained dynamic ^{19}F NMR spectra for two 19-mer modified templates during primer elongation: G*C*A* [d(5′-CTTACCATCG*CAACCATTC-3′)] and G*C*T* [d(5′-CTTACCATCG*CTACCATTC-3′)]. We found that lesion stacking is favored in the G*C*T* sequence compared to the G*C*A* counterpart. Surface plasmon resonance binding results showed consistently weaker affinities for the modified DNA with the binding strength in the order of FABP > FAF and G*C*A* > G*C*T*. Primer extension was stalled at (*n*) and near (n − 1 and *n* + 1) the lesion site, and the extent of blockage and the extension rates across the lesion were influenced by not only the DNA sequences but also the nature of the adduct's chemical structure (FAF vs FABP) and the polymerase employed (Kf-exo$^-$ vs Dpo4). Steady-state kinetics analysis with Kf-exo$^-$ revealed the most dramatic sequence and lesion effects at the lesion (*n*) and postinsertion (*n* + 1) sites, respectively. Taken together, these results provide insights into the important role of lesion-induced conformational heterogeneity in modulating translesion DNA synthesis.



A DNA adduct is DNA that is damaged by interactions of the genome with endogenously or exogenously produced reactive chemical species.[1] Although human cells are armed with various repair machineries to correct the errors,[2,3] some lesions evade the repair process and enter DNA synthesis. Replicative polymerases synthesize DNA with high accuracy and processivity. However, bulky DNA lesions can stall high fidelity replicative polymerases and often recruit specialized bypass polymerases for translesion synthesis (TLS), resulting in either error-free or error-prone DNA synthesis.[2,3] DNA lesions can cause complex mutations, depending on the structure, the nucleotide bases surrounding the lesions, their interactions with the amino acids of the polymerase, and the DNA polymerases involved during the replication process.[4,5]

Arylamines and related chemicals have been implicated in the etiology of various sporadic human cancers, including breast, liver, and bladder cancers,[1] and form C8-substituted dG lesions *in vivo*. *N*-(2′-Deoxyguanosin-8-yl)-2-aminofluorene (AF), *N*-acetyl-(2′-deoxyguanosin-8-yl)-2-aminofluorene (AAF), and *N*-(2′-deoxyguanosin-8-yl)-4-aminobiphenyl (ABP) are three extensively studied arylamine-DNA adducts (Figure 1A), which are caused by the liver carcinogen 2-aminofluorene or the well-known human bladder carcinogen 4-aminobiphenyl.[6,7] *In vitro*, AF is processed by the high-fidelity polymerases but slows down replication, whereas the bulky *N*-acetylated AAF blocks the replication process and needs bypass polymerases for TLS. Thus, AF induces point mutations, and AAF produces point and frameshift mutations.[8] In the DNA duplex, the N-deacetylated AF- and ABP-modified dG adducts tend to exist in an equilibrium of two prototype conformers: B-type, in which the carcinogen occupies the major groove of a double helical DNA without inducing any significant perturbations, and stacked (S), in which the carcinogen is intercalated between neighboring nucleobases in the duplex and the glycosidic linkage to the modified guanine is *syn* (Figure 1B).[9−13] The bulky N-acetylated AAF could additionally exist in a wedge (W) conformer, in which the hydrophobic fluorene moiety is well placed in the narrow minor groove area. The S/B conformeric equilibrium is dependent on the size, coplanarity, and topology of the arylamine carcinogen. For example, the AF/FAF adduct exists in an S/B mixture owing to the methylene linkage between the two aromatic rings that restricts the flexibility and makes it planar

Received: March 14, 2014
Revised: May 24, 2014
Published: May 29, 2014

   dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

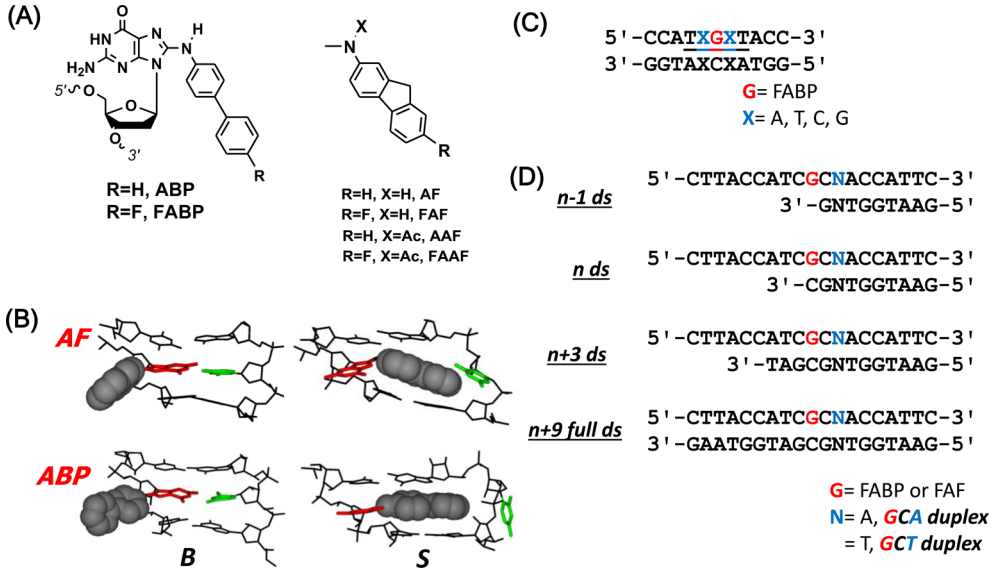


Figure 1. (A) Structures of ABP [*N*-(2′-deoxyguanosin-8-yl)-4-aminobiphenyl], AF [*N*-(2′-deoxyguanosin-8-yl)-2-aminofluorene], and AAF [*N*-(2′-deoxyguanosin-8-yl)-2-acetylaminofluorene] and their fluoro models, FABP [*N*-(2′-deoxyguanosin-8-yl)-4′-fluoro-4-aminobiphenyl], FAF [*N*-(2′-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene], and FAAF [*N*-(2′-deoxyguanosin-8-yl)-7-fluoro-2-acetylaminofluorene]. (B) Major groove views of the B and S conformers of AF and ABP (color code: modified-dG, red; dC opposite the lesion site, green; and fluorine, gray CPK); (C) The 11-mer duplexes used in the present study. (D) Duplexes used in simulated translesion synthesis.

with efficient stacking ability (Figure 1A).[14] Similarly, the more planar bulky amine adducts such as 1-aminopyrene[15]and isoquinoline adopt predominantly S- and/or W-type conformers.[16,17] In contrast, the flexible one-ring aniline and two-ring ABP adducts primarily adopt major groove B conformation.[10,18,19] As shown in Figure 1A and B, ABP differs from AF in that it lacks a methylene bridge so that the biphenyl moiety is twisted. In addition, the nucleotide sequences surrounding the lesion also dictate the lesion-induced conformational heterogeneity,[9,11,12,20−22] thus modulating repair and replication outcomes.[11,23−26] Using the fluorinated analogues of AF, AAF, and ABP (e.g., FAF, FAAF, and FABP; Figure 1A), we have shown that these arylamine lesions undergo conformation-specific nucleotide excision repair (NER), i.e., the more thermodynamically unstable S-conformational lesions are repaired more efficiently than B-conformational lesions.[26]

The mutagenicity of an adduct is also affected by its location within a DNA template and the neighboring bases. We recently studied the active site conformation of FAF in the presence of DNA polymerase β using ^{19}F NMR, nucleotide insertion assays, and surface plasmon resonance (SPR). FAF in a single nucleotide gap adopted both S and B conformations, and heterogeneity was retained upon binding to the polymerase; however, it was altered by the incoming dNTP.[27] One of the most striking examples of sequence effects involves the *Nar*I sequence (5′--$G_1G_2CG_3CN$--3′), in which AF and AAF adducts have been shown to yield higher frequencies of frameshift mutations when they are associated with the third guanine (G_3) of the sequence in *E. coli*.[28] Interestingly, mutational frequencies are affected by the nature of the base at the 3′-next flanking position (*N*); the presence of dC at the *N* position resulted in a high rate of mutation compared to that of thymidine.[29] In addition, structural studies have indicated that AF adducts adopt the S conformation when N = C but display heterogeneity when C is replaced by T.[11]

Recently, we identified an unusual 3′-flanking T effect on a random 11-mer duplex sequence (5′-CCATCG*CNACC-3′; N = T or A).[30] As shown in Figure 2A, FABP modification exhibited

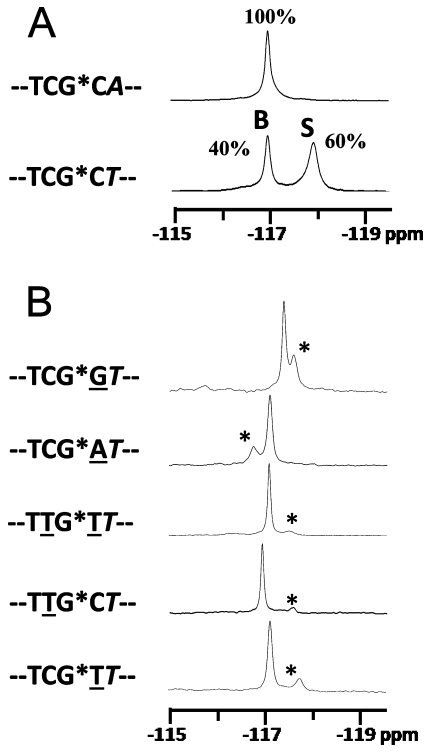


Figure 2. ^{19}F NMR of FABP-modified 11-mer duplexes. (A) Original G*CT and G*CA sequences and (B) sequences studied for the flanking base effect (see Figure 1C) at 20 °C. The 5-mer in the middle represents the core of the 11-mer duplex with the 3′-next flanking base in italics; the underlined base indicates the change compared to the original 11-mer G*CT sequence (5′-CCATCGCTACC-3′).

a 4:6 ratio of B:S conformations in the G*C*T* sequence context. When the 3′-next flanking base was changed from T to A (5′-CCATCG*CAACC-3′), the B conformation was adopted exclusively. Similar sequence effects also have been observed for FAF- and FAAF-modified G*C*N* 11-mer sequences. FAF-

dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

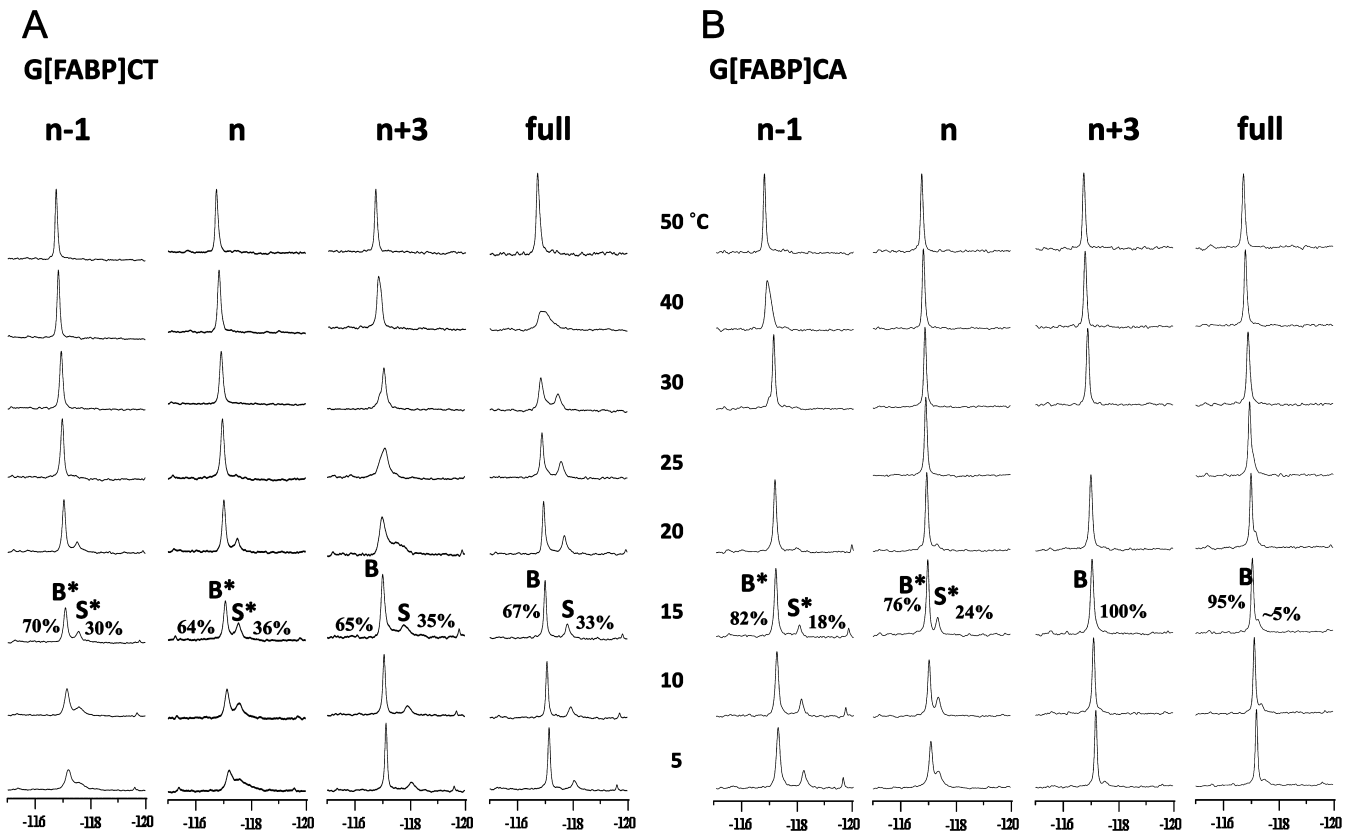


Figure 3. Dynamic ^{19}F NMR spectra of FABP-modified 19-mer (A) G*CT and (B) G*CA duplexes at different single/double strand junction positions, i.e., n − 1, n, n + 3, and full (see Figure 1D).

modified G*CA and G*CT sequences adopted the S conformation in 66% and 90% of populations, respectively. Furthermore, *E. coli* repair[30] studies have revealed that the bulky N-acetylated FAAF was repaired 3- to 4-fold more efficiently than the N-deacetylated FABP and FAF analogues, and all three adducts were repaired more efficiently in the G*C*A* sequence compared to that in the G*C*T* sequence context.

In the present study, we performed systematic dynamic ^{19}F NMR, SPR, and primer elongation kinetic studies for the TLS of the G*C*T* and G*C*A* sequences containing FAF and FABP adducts (Figure 1). The results showed that the bulky lesions on the G*C*T* duplex exhibited greater populations of the stacked S conformation compared to the G*C*A* counterparts and that the S conformation decreased the binding affinity of complementary strands. Moreover, full length primer extension experiments were performed to investigate the similar sequence effects on *in vitro* translesion synthesis by two different polymerases: high fidelity replicative polymerase exofree Klenow fragment (Kf-exo$^-$) and error prone specialized bypass polymerase Dpo4. We found that the full-length primer extensions across the modified templates were significantly faster in the G*C*A* sequence, which commonly adopted the B conformation. Significant stalling occurred due to both lesions. The *anti*-B conformation of FABP showed more stalling at the prelesion (n − 1) site, whereas the S conformation of FAF showed more stalling at the lesion (*n*) site. The relative nucleotide insertion rates were significantly reduced immediately upstream of the lesion (n + 1 and n + 3). Together, the results of this study shed light on the roles of lesion-induced conformational heterogeneity in modulating the efficiency of TLS.

MATERIALS AND METHODS

Caution: *4-Aminobiphenyl and 2-aminofluorene derivatives are mutagens and suspected human carcinogens; therefore, they must be handled with caution.*

Crude oligodeoxynucleotides (ODN, 1−10 μmol scale) in desalted form were purchased from Eurofins MWG Operon (Huntsville, AL). All HPLC solvents were purchased from Fisher Inc. (Pittsburgh, PA) and used as received. Kf-exo$^-$ and Dpo4 were received as gifts from Dr. Catherine Joyce (Yale University, CT) and Dr. F. Peter Guengerich (Vanderbilt University, TN), respectively.

Preparation of FAF- and FABP-Modified ODNs. FABP modification of 11-mer ODNs (5′--CCA*TXGXT*ACC-3′, **X** = A, T, C, G; Figure 1C) and FAF/FABP modification of 19-mer ODNs (5′-CTTACCATCG*C*N*ACCATTC-3′, G* = FABP or FAF; *N* = A or T; Figure 1D) were performed by the procedures described previously.[11,12,30,31] Briefly, 5−10 mg of *N*-acetoxy-*N*-trifluoroacetyl-7-fluoro-4-aminobiphenyl or *N*-acetoxy-*N*-trifluoroacetyl-7-fluorofluorene dissolved in absolute ethanol was added to a sodium citrate buffer (pH 6.0) containing 200−250 ODs of unmodified ODN and placed in a 37 °C shaker overnight. The modified strands were purified by a reverse phase-HPLC system which consisted of a Hitachi EZChrom Elite HPLC unit with an L2450 diode array detector and a Phenomenex Luna C18 column (150 × 10 mm, 5.0 μm). We employed a gradient system involving 3−15% acetonitrile for 25 min in ammonium acetate buffer (100 mM, pH 7.0) with a flow rate of 2.0 mL/min.

Dynamic ^{19}F NMR. Approximately 15 ODs of a modified strand was annealed with an equimolar amount of a complementary strand to produce different duplexes (Figure

dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

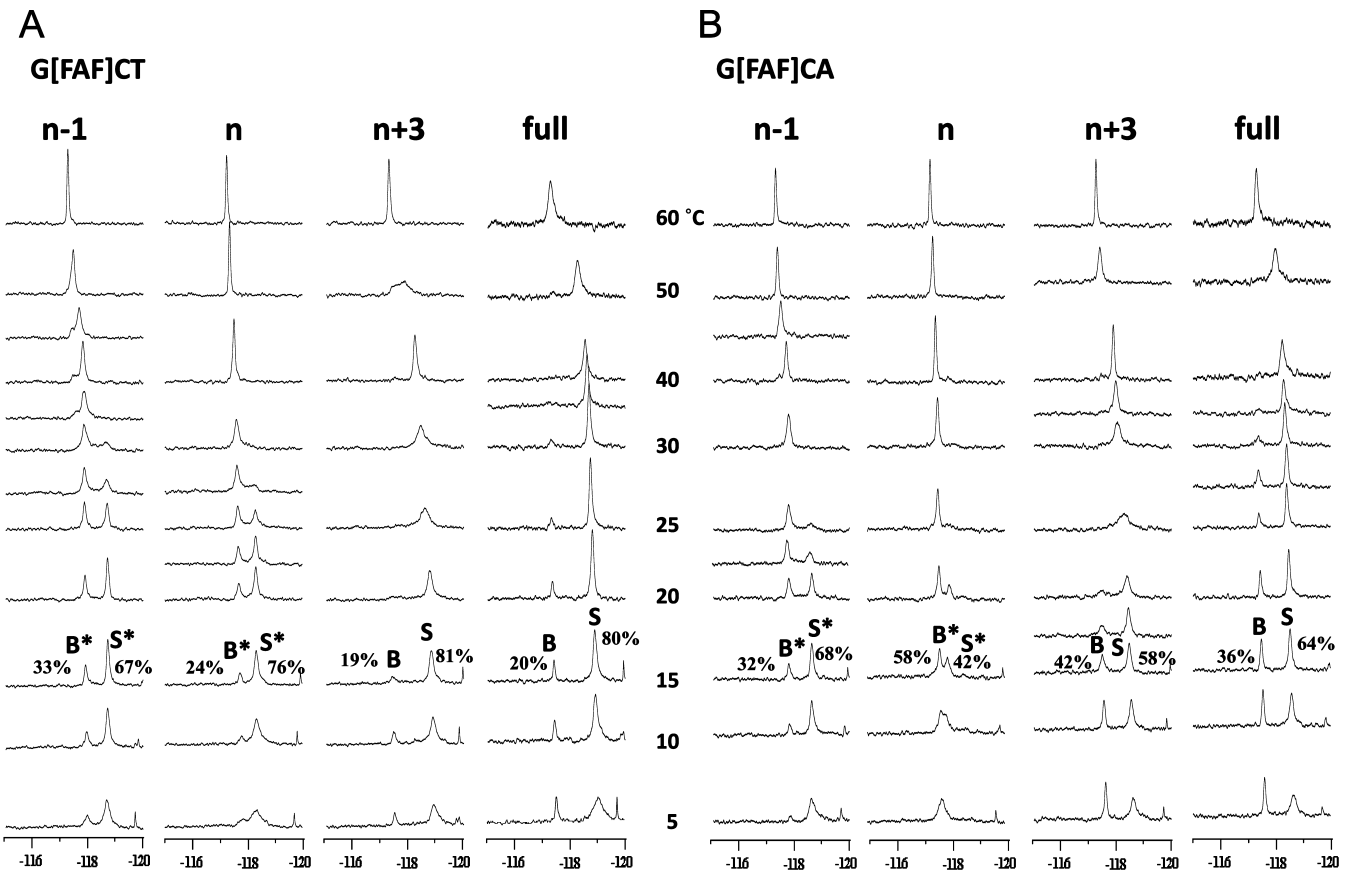


Figure 4. Dynamic ^{19}F NMR spectra of FAF-modified 19-mer (A) G*CT and (B) G*CA duplexes at different single/double strand junction positions, i.e., n − 1, n, n + 3, and full (see Figure 1D).

1C and D) that were lyophilized. The samples were then dissolved in 300 μL of typical pH 7.0 NMR buffer containing 10% D_2O/90% H_2O with 100 mM NaCl, 10 mM sodium phosphate, and 100 μM EDTA, and filtered into a Shigemi tube through a 0.2 μm membrane filter. All ^1H and ^{19}F NMR results were recorded using a dedicated 5 mm $^{19}F/^1H$ dual probe on a Bruker DPX400 Avance spectrometer operating at 400.0 and 376.5 MHz, respectively, using acquisition parameters described previously.[11,12,32,33] Imino proton spectra (Figure S1, Supporting Information) at 5 °C were obtained using a phase-sensitive jump-return sequence and referenced relative to that of DSS. ^{19}F NMR spectra were acquired in the ^1H-decoupled mode and referenced relative to that of $CFCl_3$ by assigning external C_6F_6 in C_6D_6 at −164.9 ppm. Dynamic ^{19}F NMR spectra (Figures 2, 3, and 4) were measured between 5 and 60 °C with an increment of 5−10 °C. Temperatures were maintained by a Bruker-VT unit with the aid of controlled boiling of liquid N_2 in the probe.

Primer Extension Assays. Steady-state kinetic experiments were performed as described previously.[34,35] Briefly, the primers (29−33-mers) were 5′-radiolabeled using [γ-^{32}P]ATP and T4 polynucleotide kinase (T4 PNK) 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 heating to 95 °C for 5 min and then slowly cooled to room temperature in 3 h. The primer−template sequence (100 nM) was incubated with Kf-exo$^-$ (0.5 or 1.0 nM) for 5 min to form a binary complex in Tris buffer (Tris, 50 mM at pH 7.4; BSA, 50 μg/mL; 5% (v/v) glycerol). The reaction was initiated by adding a dNTP/$MgCl_2$ (dNTP in variable concentration and 5 mM $MgCl_2$) solution to a binary mixture and incubated at 20 °C. The reaction was quenched at different time intervals using quenching buffer (containing 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 2,000 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.

Running Start Experiments. A 25-mer primer was used with the FAF- or FABP-modified dG at position 31 of the 44-mer template (Figure 5). The extension was performed using both Kf-exo$^-$ (2.5 nM) and Dpo4 (50 nM FABP and 100 nM FAF) polymerase in the presence of all four dNTPs. Aliquots were withdrawn at regular intervals of time, and the reaction was quenched as mentioned above. The extended products were separated on a 20% denaturing polyacrylamide gel.

Steady-State Kinetics Experiments. Steady-state kinetic parameters for incorporation of the nucleotide opposite the unmodified and FAF- or FABP-modified templates were determined by following the procedures reported previously.[34,35] The reactions were performed at 20 °C. For the unmodified sequence, reactions were performed in a shorter time period of 1 min for correct nucleotide incorporation and for up to 45 min in the case of 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). The kinetic parameters k_{cat} and K_m were determined as described earlier.[34,35]

dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

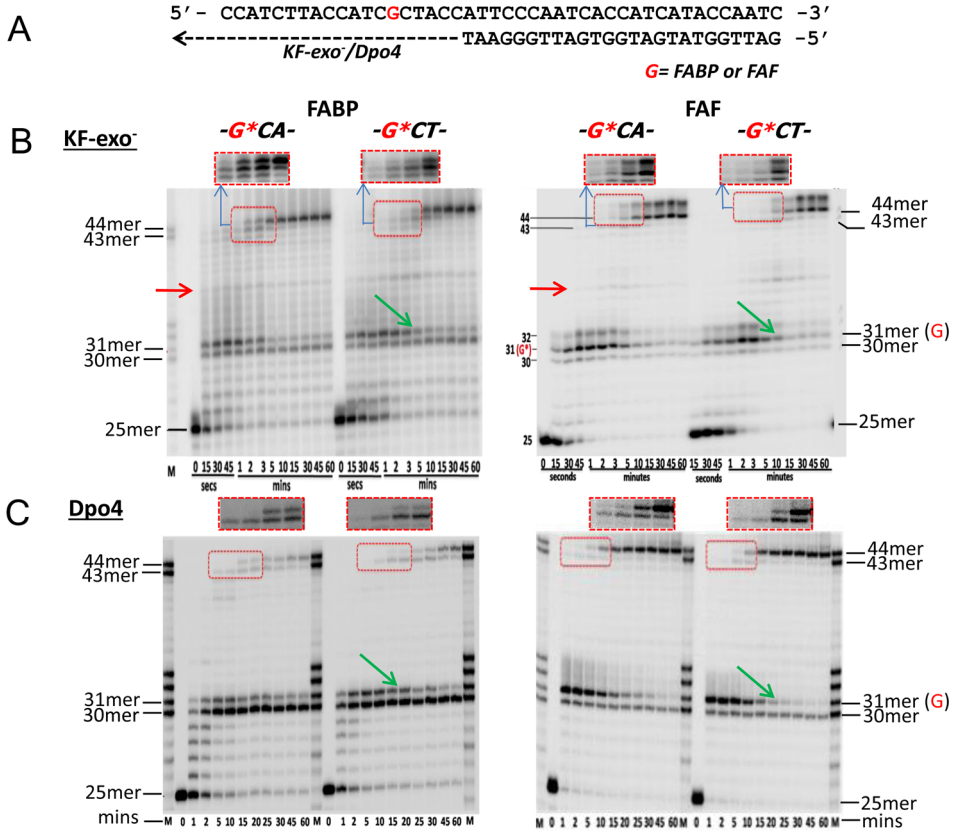


Figure 5. (A) Template and primer sequences used for running start experiments. Gel results across FABP (left)- and FAF (right)-modified templates (G*CA and G*CT) in the presence of all four dNTPs (250 μM each) with (B) Kf-exo$^-$ (2.5 nM) or (C) Dpo4 (50 nM, FABP; 100 nM, FAF) at different time intervals (M, marker).

Surface Plasmon Resonance (SPR). SPR measurements were conducted using a Biacore T200 instrument (GE Healthcare) in order to probe the effect of lesion-induced conformational heterogeneity using different lengths of complementary strands from n − 1 to n + 5 full duplex in a polymerase-free condition (Figure 7A). A 14-mer biotinylated DNA strand (5′-biotin-CTATCG*CN*ACCATC-3′, *N* = T or A) was used for the SPR work. The template strands have the same sequence contexts as those utilized in the NMR studies (Figure 1D) except for the variation in the underlined portions (CC to CT; ATT to ATC) to avoid complications in complementary strand elongation (n − 1, 5′-GATGGTXG-3′; n, 5′-GATGGTXGC-3′; n + 1, 5′-GATGGTXGCG-3′; n + 2, 5′-GATGGTXGCGA-3′; n + 3, 5′-GATGGTXGCGAT-3′; and n + 5, 5′-GATGGTXGCGATAG-3′; X = T for the GCA sequence, and X = A for GCT sequence; see Figure 7A). The biotinylated oligonucleotides were modified with FABP and FAF by using the usual biomimetic procedure, purified by HPLC and characterized using MALDI-TOF mass spectrometry (Supporting Information, Figures S2 and S3). Streptavidin (SA) was immobilized on a CM5 chip using the EDC-NHS coupling method using the manufacturer's protocol. The unbound SA was removed from the chip by injecting five 1 min pulses of NaOH (50 mM), and we stabilized the surface with running buffer for 30 min. The unmodified and modified oligo strands were coated over the SA surface to about 450 resonance units (RU) through a manual command mode. The stability of duplexes was monitored by injecting different concentrations of complementary strands (analytes) in HBS-P+ buffer (10 mM Hepes at pH 7.4; 150 mM NaCl; and 0.05% surfactant P20) at 25 °C. Dissociation rate constant (k_d) for each duplex were determined by fitting the data in a 1:1 model using the k_d-alone fitting method available in Scrubber software, version 2.0 (Myszka and collaborators; BioLogic Software) (Supporting Information, Table S10). The concentration-independent parameter k_d was calculated by fitting the initial 100 s of the dissociation curves (Supporting Information, Figure S4). The goodness of the fit was determined from the residual standard deviation.

RESULTS

Experimental DNA Sequences. The chemical structures of ABP, AF, and AAF and their corresponding fluorine analogues, FABP, FAF, and FAAF, are shown in Figure 1A. In addition, the 11-mer G[FABP]CT duplex sequence used in our previous repair studies and the flanking sequence variations used are shown in Figure 1C.[30] These modified DNA strands were annealed with complementary strands to form duplexes for the ^{19}F NMR measurements. The CG*C*N* (*N* = A or T) series designed for the TLS experiments are shown in Figure 1D. The original 11-mer oligonucleotides were extended to 19-mers to improve thermal stability for the n − 1 and n duplexes. As shown in Figure 1D, four 19-mer G*CN strands (G* = FABP or FAF: N = A or T) were each annealed with complementary strands of variable lengths (n − 1, n, n + 3, and n + 9; n is the lesion site) to create four discrete model TLS systems.

Flanking Base Effects. The rationale behind this study is to investigate the uniqueness of 5′- or 3′-flanking C in the -CG*CT- sequence context in promoting lesion-induced conformational heterogeneity. As such, the 3′- and 5′-flanking C of CG[FABP]CN (*N* = A or T) was systematically switched to A, T, or G

dx.doi.org/10.1021/bi5003212 | Biochemistry 2014, 53, 4059−4071

(Figure 1C). FABP was chosen over FAF because of its dramatic sequence-dependent S/B-conformational heterogeneity (Figure 2A).[30] The ^{19}F NMR results at 20 °C are shown in Figure 2B and were compared with the original FABP-modified -CG*CA- and -CG*CT- duplexes (Figure 2A). The full-range ^{1}H imino and dynamic ^{19}F NMR spectra of the modified duplexes are shown in Supporting Information, Figure S1.

In all modified cases (Figure 2B), one major ^{19}F signal was observed at approximately −117 ppm. This result is consistent with the typical chemical shift range for the FABP-modified B-conformer.[31] These duplexes also exhibited a small intensity signal (*) in the upfield region (−117.5 to −118.0 ppm). Assignment of these minor signals was difficult, but they could be either the stacked S- or minor groove W-type conformers as their fluorine atoms are known to be upfield shifted relative to the external binding B-type conformer (see below for more ^{19}F NMR assignments).[14,21,33] The G*AT sequence was an exception and showed a small downfield signal (∼116.8 ppm), which could be regarded as a variation of the B-type conformer. These results indicated that the T adjacent to the lesion site promoted the B-type conformation, while replacement of the 3′-flanking C with A or G increased the conformational heterogeneity; however, the effect was not as great as that of the 3′-next flanking T (Figure 2A). Therefore, it was concluded that the CG*CT sequence context is very unique in promoting the S conformation.

Conformational Heterogeneity at Simulated TLS Single/Double Strand Junctions. The goal of this study was to determine whether the S/B-conformational heterogeneity in the G*CA and G*CT duplexes also presents at the single/double strand junctions formed during the primer elongation in TLS process. Therefore, we conducted ^{19}F NMR experiments in buffer-only solutions in the absence of a polymerase on four discrete 19-mer G*CN duplexes (G* = FABP or FAF: N = A or T) (Figure 1D). These four model TLS systems were prepared by annealing the 19-mer modified strand with complementary strands of variable lengths (n − 1, n, n + 3, and n + 9; n is the lesion site).

FABP. Figure 3A−B shows the dynamic ^{19}F NMR spectra (5−50 °C) for TLS of the FABP-modified G*CT and G*CA sequences (Figure 1D) at various elongation positions with a temperature range between 5−50 °C. Signals were assigned according to published procedures based on relative chemical shifts, dynamic NMR signal patterns, and H/D isotope effects.[33,35] ^{19}F shielding is a hallmark of the van der Waals interactions and the ring current effects caused by the carcinogen moiety within the stacked and bulge duplexes (S-type conformation). We have studied a number of ^{19}F NMR spectra of arylamine-modified duplexes (including FABP and FAF) in various sequence settings and in all cases found that the fluorine of the S-type conformer resonates upfield relative to that of the external binding B-type conformer.[12,33] The same trend has been observed with various fully paired and deletion duplexes modified with the bulky N-acetylated FAAF, which specifically revealed a mixture of B, S, and W conformations in the −115.0 to −115.5 ppm, −115.5 to −117.0 ppm, and −116.5 to −118.0 ppm ranges, respectively.[21] In contrast, the N-deacetylated FABP and FAF adducts adopted an interchangeable mixture of the B- and S-conformers. Accordingly, the signals in Figures 3 and 4 were assigned as the B (downfield)- and S (upfield)-type conformers, respectively.

The fully paired 19-mer G*CT and G*CA duplexes (Figure 1D), which represent the end points of TLS, displayed comparable conformational differences (Figure 3A−B) similar to those of their respective 11-mer duplexes (Figure 2A). Populations of G*CT sequences in the S conformation decreased significantly from the 11-mer (60%) to the 19-mer (33%) duplexes. The ^{19}F signal patterns for the n + 3 and full duplex sequences for both G*CA and G*CT did not change much. The ^{19}F NMR signals at the n and n − 1 sites were designated as B-like (B*) and S-like (S*), respectively, based on their chemical shift similarities to the duplexes. The B (B*)- and S (S*)-proportions were maintained as elongation progressed from n − 1 to full in the G*CT TLS series. Interestingly, we observed a significantly larger population of the S conformation at the n − 1 and n positions as compared to that in the n + 3 or n + 6 position in the G*CA series (Figure 3B). The conformer population ratios at different positions for each modified strand are summarized in Table 1.

Table 1. Summary of the Conformational Profile Exhibited by FABP and FAF at Different Complementary Strand Positions in the G*CA and G*CT Sequence Context at 15 °C

		conformational profile			
duplex	sequence context	FABP		FAF	
n − 1	G*CA	B* (82%)	S* (18%)	B* (32%)	S* (68%)
	G*CT	B* (70%)	S* (30%)	B* (33%)	S* (67%)
n	G*CA	B* (76%)	S* (24%)	B* (58%)	S* (42%)
	G*CT	B* (64%)	S* (36%)	B* (24%)	S* (76%)
n + 3	G*CA	B (100%)		B (42%)	S (58%)
	G*CT	B (65%)	S (35%)	B (19%)	S (81%)
full	G*CA	B (95%)	B* (5%)	B (36%)	S (64%)
	G*CT	B (67%)	S (33%)	B (20%)	S (80%)

FAF. The dynamic ^{19}F NMR spectra for the FAF-modified G*CT and G*CA TLS series (n − 1 to full) are shown in Figure 4A−B. In both sequences, FAF exhibited B (B*)- and S (S*)-conformation heterogeneity, and G*CT had approximately 16% more S-conformers than G*CA in the full duplex. Like the FABP adduct, template strand elongation from the 11-mer to the 19-mer duplex reduced the G*CT S-conformer population from 90% to 80%. At n + 3, the S-conformer population differed by approximately 23% between the two sequences, primarily due to the increase in the B-conformer population in the G*CA template (Figure 4B). Moving from the n + 3 to the n and the n − 1 positions, FAF exhibited two ^{19}F signals. The S conformation gap was further increased to 34% at the lesion site (n), mostly due to the increase in the B-conformer population of G*CA (Figure 4B and Table 1). However, the n − 1 series displayed very similar ^{19}F characteristics and a 3:7 ratio of B- and S-conformations. The conformational results of the FAF-induced TLS system are summarized in Table 1.

Primer Extension Experiments. The purpose of primer extension experiments was to determine the impact of lesion-induced conformational heterogeneity on the polymerase efficiency during TLS. Therefore, experiments were performed using 44-mer G*CN (G* = FABP or FAF; N = A or T) templates. As shown in Figure 5A, ^{32}P-labeled 25-mer primers were annealed to the templates, and primer extension was carried out in the presence of all four dNTPs and polymerases (Kf-exo$^-$ or Dpo4). Running start experiments were specifically carried out to determine the overall impact on the rate of full length primer extension across the lesion and the major blockage sites for polymerase.

dx.doi.org/10.1021/bi5003212 | Biochemistry 2014, 53, 4059−4071

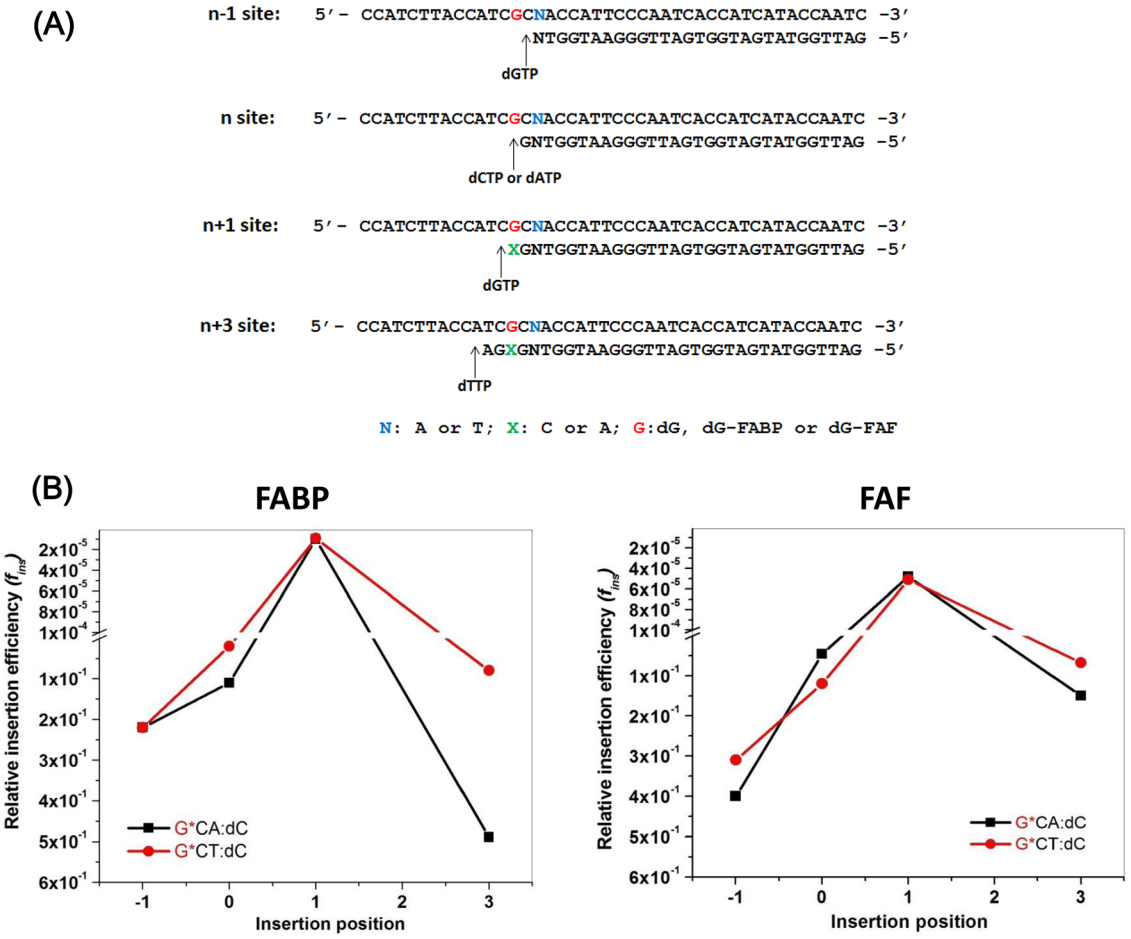


Figure 6. (A) Oligonucleotide sequences used for steady state kinetics. (B) Plots of the extension frequency (f_{ins}) vs insertion site for FABP- and FAF-modified dC match series.

Running Start Experiments on FABP Adducts with Kf-exo⁻ Polymerase. The products of Kf-exo⁻-mediated primer extension across FABP at different time intervals at room temperature are shown in Figure 5B (left). Kf-exo⁻ was able to extend the primer to a full-length 44-mer across G*C*T* and G*C*A* in 2−5 min. However, two major stalls in extension were observed: one at the nucleotide before the lesion site (n − 1, 30-mer) and the other at the lesion site (n). The blockage at n was stronger than that at n − 1. The full-length product was formed from this blocked primer, although the primers at n − 1 persisted even after incubation for 60 min. Additionally, both sequences exhibited unusual blockage at the site five bases upstream of the lesion (red arrows). Although the sequences G*C*A* and G*C*T* both showed similar blockage patterns, the durations of the blockages were different. In the initial few minutes, G*C*T* displayed a significant increase in blockage at the lesion site (green arrow), which decreased after the full-length product was produced. A similar pattern was observed for G*C*A*, but TLS occurred in a shorter time (see expanded red rectangular insets in Figure 5B). The prolonged blockage at the lesion site (n) was converted to a full-length product with G*C*A* displaying a faster rate of extension than G*C*T*.

Running Start Experiments of FABP Adducts with Dpo4 Polymerase. Similar running start experiments were performed with Dpo4 in order to compare the lesion and sequence effects between a replicative polymerase and a bypass polymerase (Figure 5C, left). Dpo4 extended the primer to a full-length 44-mer and displayed major blockages at n − 1 and n. Compared to Kf-exo⁻, the extent of full-length extension was significantly less, even at a high enzyme concentration (100 nM, data not shown). In addition, a stronger blockage occurred at n − 1 with Dpo4, compared with that using Kf-exo⁻. As with Kf-exo⁻, the G*CT sequence exhibited an extended blockage at n (green arrow), resulting in slower extension of the primer to full length (rectangular insets, Figure 5C).

Running Start Experiments of FAF Adducts. The results of the running start experiments for the FAF adducts are shown in Figure 5B (right). Here, the primers were blocked at n and, to a lesser extent, at n − 1 and n + 1. Regardless, Kf-exo⁻ was able to extend more efficiently across the FAF adduct compared to that across FABP. In addition, a blunt-end addition was also observed. A similar sequence effect was observed with the FAF adduct, and G*CA showed a faster rate of extension compared to that of G*CT (see expanded red rectangular insets, Figure 5B). As shown in Figure 5C (right), primer extension was stalled at two sites, n − 1 and n, in the Dpo4-mediated primer extension. Unlike with Kf-exo⁻, there was no blockage at n + 1; however, a slightly greater rate of extension across G*C*A* compared to that across G*C*T* was observed.

Steady-State Kinetics. Steady-state kinetic experiments in the presence of Kf-exo⁻ were performed in order to investigate the roles of lesion-induced conformational heterogeneity in nucleotide insertion kinetics. As shown in Figure 6A, lesions were positioned one nucleotide downstream of the template base (n −

dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

1), at the lesion position (n), one nucleotide upstream (n + 1), or three nucleotides upstream (n + 3). We examined the lesion effects using the relative insertion efficiency f_{ins}, which was defined as $(k_{cat}/K_m)_{modified\ or\ mismatched}/(k_{cat}/K_m)_{unmodified}$. Figure 6B shows plots of the f_{ins} as a function of nucleotide insertion position. The results are summarized in Table 2, and further details including standard deviations are provided in Supporting Information, Tables S1–S9.

Table 2. Summary of the Insertion Efficiency against the FABP- and FAF-Modified Template in the Presence of Kf-exo$^-$ at Different Insertion Sites in the G*C*A* and G*C*T* Sequence Context[a]

			insertion efficiency (f_{ins})[b]	
insertion site	sequence context		FABP	FAF
n − 1	-G*CA-	dGTP[c]	2.2×10^{-1}	4.0×10^{-1}
	-G*CT-	dGTP[c]	2.2×10^{-1}	3.1×10^{-1}
n	-G*CA-	dCTP	1.1×10^{-1}	4.6×10^{-2}
		dATP	3.0×10^{-4}	1.5×10^{-3}
	-G*CT-	dCTP	2.0×10^{-2}	1.2×10^{-1}
		dATP	1.8×10^{-4}	1.5×10^{-2}
n + 1	-G*CA-	dG*:dC[d]	9.6×10^{-6}	4.8×10^{-5}
		dG*:dA[d]	ND	8.5×10^{-7}
	-G*CT-	dG*:dC[d]	9.1×10^{-6}	5.1×10^{-5}
		dG*:dA[d]	ND	2.1×10^{-7}
n + 3	-G*CA	dG*:dC[e]	4.9×10^{-1}	1.5×10^{-1}
		dG*:dA[e]	2.0×10^{-2}	1.3×10^{-2}
	-G*CT-	dG*:dC[e]	8.0×10^{-2}	6.8×10^{-2}
		dG*:dA[e]	2.0×10^{-3}	4.3×10^{-3}

[a]Detailed kinetic analysis with error limits are provided in the Supporting Information (Tables S1 to S8). [b]The relative insertion efficiency $f_{ins} = (k_{cat}/K_m)_{modified\ or\ mismatch}/(k_{cat}/K_m)_{unmodified}$. [c]Incoming nucleotide. [d]Template–primer terminus. [e]C or A opposite G*.

Insertion Efficiency. At the prelesion site n − 1, the lesion effect was minimal for the insertion of the correct nucleotide dGTP (for the G*CA and G*CT sequences, FABP, $f_{ins} = 2.2 \times 10^{-1}$ and 2.2×10^{-1}; FAF, $f_{ins} = 4.0 \times 10^{-1}$ and 3.1×10^{-1}, respectively; Figure 6B and Table 2). The f_{ins} values for the correct dCTP opposite FABP at the lesion site (n) were reduced to 1.1×10^{-1} and 2.0×10^{-2} for G*C*A* and G*C*T*, respectively, which amount to 10- and 50-fold reductions relative to the controls. The f_{ins} for the wrong nucleotide dATP was reduced significantly in the G*C*A* (3.0×10^{-4}) and G*C*T* (1.8×10^{-4}) contexts. These results suggested that the insertion of dCTP was preferred over dATP by 367- and 111-fold in the G*C*A* and G*C*T* sequences, respectively (Table 2). The f_{ins} values of dCTP opposite FAF were reduced to 4.6×10^{-2} and 1.2×10^{-1} for G*C*A* and G*C*T*, respectively, representing rate reductions of 22- and 8-fold, respectively. The f_{ins} of the incorrect dATP opposite the lesion in the G*C*A* sequence was 1.5×10^{-3}, and in the G*C*T* sequence, it was 1.5×10^{-2}, indicating that the dCTP insertion was preferred over the dATP insertion by 31-fold in G*C*A* and 8-fold in G*C*T* (Table 2).

For the FABP adducts, the f_{ins} values of dGTP opposite C at the n + 1 position were 9.6×10^{-6} and 9.1×10^{-6} in the G*C*A* and G*C*T* sequences, respectively. However, when the incorrect dA was present opposite the lesion, nucleotide insertion was blocked for both sequences and the efficiency could not be determined (ND, Table 2). For the FAF adducts (Figure 6B), the efficiencies of the correct nucleotide dGTP at n + 1 were similar in both the G*C*A* (4.8×10^{-5}) and G*C*T* (5.1×10^{-5}) sequences. The f_{ins} values for dGTP at n + 1 were reduced by 6- and 242-fold in the G*C*A* and G*C*T* sequences, respectively, when the incorrect dA was paired at the lesion site. However, the effect of a lesion at n + 3 only minimally affected the efficiency. The f_{ins} values for the FABP adducts were 4.9×10^{-1} (G*C*A*) and 8.0×10^{-2} (G*C*T*), and the f_{ins} values for the FAF adducts were 1.5×10^{-1} (G*C*A*) and 6.8×10^{-2} (G*C*T*) for the correct insertions (Table 2). For dA opposite the lesion site, the f_{ins} values at n + 3 between the G*C*A* and G*C*T* sequences differed by 10- and 3-fold in the FABP and FAF adducts, respectively.

SPR Binding Affinity Measurements for Simulated TLS Single/Double Strand Junction Duplexes. The objective of SPR experiments was to measure real-time association between

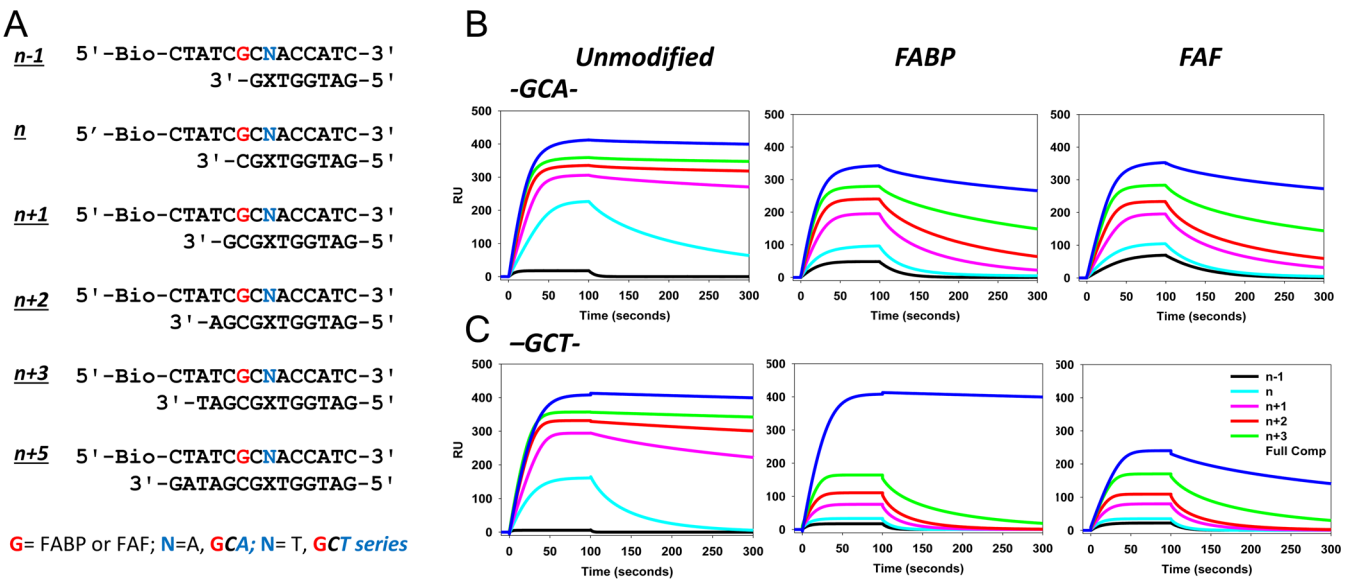


Figure 7. (A) Primer–template duplex sequences for SPR measurements. SPR binding responses of complementary strands opposite unmodified and modified DNA for (B) GCA and (C) GCT sequence contexts. Strands concentrations are 50 nM (n − 1), 100 nM (n), 150 nM (n + 1), 200 nM (n + 2), and 250 nM (n + 3 and n + 5 full length complementary strands).

dx.doi.org/10.1021/bi5003212 | Biochemistry 2014, 53, 4059–4071

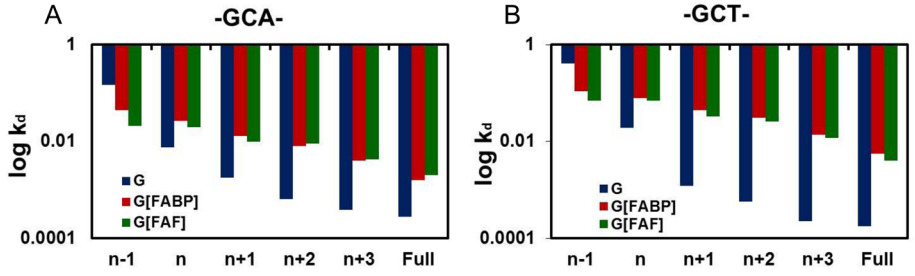


Figure 8. Plot of SPR log k_d vs complementary strand length: (A) GCA and (B) GCT sequence contexts. See Figure 7A for sequences.

template and complementary strands in the absence of a polymerase. Binding data were obtained by injecting different lengths (8-mer n − 1, 9-mer n, 10-mer n + 1, 11-mer n + 2, 12-mer n + 3, and 14-mer n + 5) of complementary strands to lesions containing 14-mer template strands (GC*T* vs GC*A*) coated on a streptavidin chip. For each elongation, it was necessary to employ different concentrations of complementary strands to achieve steady-state associations (Figure 7: 50 nM (n − 1), 100 nM (n), 150 nM (n + 1), 200 nM (n + 2), and 250 nM (n + 3 and n + 5). Following each binding phase, the template-containing chip was washed with buffer as described in Materials and Methods. Figure 7B shows the representative sensograms as a function of different lengths of complementary strands. As a result, the usual K_D binding (k_d/k_a) kinetics of duplex formation for each elongation was not obtained. Since all of the experimental conditions included are identical for each elongation, direct qualitative comparison of RU values for each elongation is presented. Similarly, dissociation rates (k_d) are concentration-independent and thus could be used to estimate binding strengths for different lengths of complementary strands. Supporting Information, Figure S4 shows curve data analysis using Scrubber (BioLogic Software). Supporting Information, Table S10 lists the respective dissociation rate constants. The results are also summarized graphically in Figure 8 as plots of log k_d vs various elongation points for GCA and GCT sequences.

FABP. SPR data shows generally greater binding activities (RU) (Figure 7B) and k_d (Figure 8) values as the complementary strand length increased from n − 1 to a full duplex in both sequences, with increasing duplex stability. Overall, we observed little difference between the unmodified GCA and GCT sequences. The effect of a lesion modification was evident with modified duplexes, which showed a substantial reduction in RU responses for the n to n + 3 sites (Figure 7B). Interestingly, however, the FABP modification caused an increase of RU values at the prelesion n − 1 site compared to that in the unmodified controls. These results indicate a lesion-induced stabilization which is in agreement with the slower dissociation at the prelesion site (Supporting Information, Table S10 and Figure 8). The effect of a lesion modification was evident with faster dissociations (k_d) than the unmodified one at the n to n + 5 site (Figure 8 and Supporting Information, Figure S10). For example, the binding strength for G[FABP]C*T* at n + 1 was decreased 3.3 (0.0443/0.0113)- and 37 (0.0443/0.0012)-fold relative to the G*C*C*A* and the unmodified GCT sequences, respectively (Supporting Information, Table S10). The sequence effect was found to be consistently around 3−4-fold different between the G*C*C*T* and G*C*C*A* sequences, respectively: n + 2 (3.8-fold, 0.0307/0.00802), n + 3 (3.7-fold, 0.00142/0.00378), and n + 5 (3.6-fold, 0.00571/0.00156).

FAF. A similar lesion and sequence effects were observed for the FAF adducts. The RU intensities of all the elongations except for n − 1 were generally suppressed compared to those of the unmodified controls (Figure 7). For example, the binding strength for G[FAF]C*T* at n + 1 was decreased 2.9 (0.0326/0.012)- and 27.1 (0.0326/0.0012)-fold relative to the G*C*C*A* and the unmodified GCT sequences, respectively (Supporting Information, Table S10). As in the case of the FABP adducts, the sequence effect was maintained at 2−3-fold differences between the G*C*C*T* and G*C*C*A* sequences.

DISCUSSION

We studied a range of structural, biochemical, and biophysical properties of a pair of arylamine-DNA lesions, one (ABP) derived from the very important human bladder carcinogen 4-aminobiphenyl that is present in tobacco smoke and a second (AF) stemming from the model carcinogen 2-aminofluorene. These two lesions, bound to the C8-position of dG, differ only in that the biphenyl moiety lacks a methylene bridge that restrains the aminofluorene ring to planarity so that the biphenyl is twisted. Specifically, we characterized lesion-induced TLS involving two unique G*C*C*A* and G*C*C*T* sequences by dynamic ^{19}F NMR, primer extension studies with two representative polymerases, Kf-exo$^-$ and Dpo4, with detailed steady state kinetic parameters, and surface plasmon resonance measurements.

3′-Next Flanking Base Effect on Lesion-Induced S/B Heterogeneity during Simulated TLS. The fully paired 19-mer G[FABP]C*T* duplex exhibited a 67:33 ratio of B- and S-conformers (Figure 3A and Table 1), consistent with the 60:40 B:S ratio that was previously observed for the 11-mer counterparts.[30] A similar conformer ratio was intact at the n − 1 and n sites. In contrast, the G[FABP]C*A* duplex adopted the B conformation predominantly (95% B:5% S) (see n + 6, Figure 3B). The FAF-modified 19-mer duplex (Figure 4) exhibited S/B heterogeneity comparable to that of the previously studied 11-mer:[30] 90% S vs 80% S for G*C*C*T* and 66% S vs 64% S for G*C*C*A*. Similar conformational heterogeneity persisted even at the n − 1 and n positions.

In general, the larger S-conformer population for the FAF adducts compared to the FABP adducts may be attributed to the presence of single carbon atom methylene bridging in FAF.[12,36] This linkage restricts the twisting between the two aromatic rings that is possible in FABP, thus enhancing the stacking surface (Figure 1A,B).[10,30,37] This simple topological difference is fascinating in its impact on the structural and biochemical properties of the lesion. Similar conformational behavior is observed with other planar bulky amine adducts such as 1-aminopyrene (AP) and the food-borne heterocyclic amine isoquinoline (IQ), which favor S and/or wedge (W)-type conformers.[16,17] On the contrary, the AF analogue imidazopyridine (PhIP) occupies the minor groove largely due to the freely rotatable phenyl at C6 position.[38] These results clearly signify the

dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

role of a single linkage in dictating the conformations adopted by FAF.

A large S-conformational gap was noted in G*C<u>A</u> vs G*C<u>T</u> for FABP (30% vs 18%) at n − 1, while no such difference (67% vs 68%) was observed for the FAF adduct (Figures 3−4 and Table 1). At the lesion site n, FAF in the G*C<u>T</u> sequence exhibited a 34% higher S-conformer population than G*C<u>A</u> (Table 1). Complementary strand elongation from n to n + 3 increased the S-conformer population in both G*C<u>A</u> (42−58%) and G*C<u>T</u> (76−81%) for the FAF adduct. For fully paired duplexes, the S-conformer population of G*C<u>A</u> increased to 64%, whereas the conformation of G*C<u>T</u> was largely unaffected (80%). Taken together, these results indicated that lengthening the duplex sequence from 11 to 19 nucleotides had no major effects on the S/B heterogeneity of FABP or FAF adducts in the G*C<u>A</u> or G*C<u>T</u> sequences. The unusual 3′-next flanking base effects observed in the fully paired duplex were largely maintained at various stages of complementary strand elongation, including the preinsertion (n − 1) and lesion (n) sites.

We performed SPR experiments to investigate the impact of lesion heterogeneity on complementary strand binding (Figure 7). SPR is a chip-based and label-free procedure that allows real-time monitoring of various replication and repair interactions involving DNA lesions.[27,39−43] SPR has also been used to study DNA assembly, DNA hybridization,[44,45] and triplex formation.[46] We obtained simple sensorgrams as a function of different complementary strands for each elongation, and consequently, we were unable to conduct the usual global binding analysis using K_D affinity values (see Results). Instead, we utilize RU intensities (Figure 7) and k_d dissociation rates (Supporting Information, Table S10) to assess lesion and sequence effects. In general, the binding strength of complementary strands toward the unmodified DNA sequence was higher than the modified sequences. In contrast, the effect of FABP or FAF modification was sensed primarily around the lesion (n − 1 to n + 2), but then the effect was reduced at n + 3 and full (n + 5) duplexes. It is worth noting that the RU for G*C<u>T</u> were consistently less than those for G*C<u>A</u> throughout the elongation. There was a trend in lesion-induced differences in terms of RU values in the G*C<u>A</u>/G*C<u>T</u> sequence context, i.e., G*C<u>T</u> was a DNA destabilizer compared to G*C<u>A</u>. Figure 8 shows graphical plots of log k_d dissociation rates vs various elongation points for the G*C<u>A</u> and G*C<u>T</u> TLS series. It is clear that the lesion effect is seen throughout the simulated elongation process (n to n + 5) with greater impact for the G*C<u>T</u> compared to that for G*C<u>A</u>. As shown in Figure 8, however, the nature of the lesion (FABP vs FAF) had no discernible effect on k_d values. Although no polymerases were involved, these SPR results are in line with the NMR and gel assay results as the greater S-conformation of the G*C<u>T</u> sequence posed a major hindrance toward the binding affinity. In other words, the lesion effect was propagated primarily upstream of the template strand and gradually diminished once the adduct came out of the pocket of the polymerase active site.[33,34,47]

Conformation Effects on Stalling during TLS in the Presence of Kf-exo⁻ and Dpo4. Kf-exo⁻ is a high-fidelity replicative polymerase lacking the 3′−5′ proofreading exonuclease activity used for TLS of various types of DNA damage, including bulky arylamine lesions.[34] High fidelity polymerases generally are blocked by bulky lesions.[5] The DNA polymerase Dpo4 from *Sulfolobus solfataricus* is a Y-family bypass polymerase, characterized by a loose active site that is accessible to solvent, thus known to promote TLS.[48] The running start results (Figure 5) showed that primer extension was stalled at the lesion site (n) and near the lesion site (n − 1 and n + 1). The extent of blockage was clearly influenced by the nature of adduct structure, i.e., both FABP and FAF lesions blocked Kf-exo⁻ at n − 1 and n and additionally at n + 1 for the FAF adducts. Both lesions strongly blocked at the lesion site (n); however, they exhibited similar sequence effects, i.e., the G*C<u>A</u> sequence exhibited a greater extension efficiency compared to that of G*C<u>T</u>. With Dpo4, the lesions stalled primer extension only at n − 1 and n, but the major blockage site occurred at n − 1 for FABP, compared to n for FAF. The different stalling characteristics could be due to the spacious nature of the Dpo4 active site compared to that of Kf-exo⁻. The high steric flexibility of Dpo4 has been shown to allow bulky and distorted DNA lesions to proceed through the active site differently.[49−52] Surprisingly, however, we observed that Dpo4 exhibited a significantly slower rate of primer extension for FABP, even at higher enzyme concentrations (Figure 5C, left). Dpo4 also causes deletions with bulky adducts.

The G*C<u>T</u> and G*C<u>A</u> sequences displayed similar blocking patterns (Figure 5), consistent with their conformational profiles around the lesion sites (Table 1). However, full-length extension across the G*C<u>A</u> sequence was faster compared to that across the G*C<u>T</u> sequence of both lesions as well as in the presence of Kf-exo⁻ and Dpo4 (insets, Figure 5B and C). As mentioned above, the polymerase prefers the modified dG in the *anti*-glycosyl conformation for efficient replication across the lesion. The relatively slower extension across G*C<u>T</u> may be due to the higher percentages of the S-conformer at the lesion site.

Hsu et al.[47] have reported the crystal structures of AF during accurate replication by BF, a high-fidelity DNA polymerase analogous to Kf-exo⁻. They found that the AF adduct exhibited a *syn* conformation at the preinsertion site (n − 1) and underwent a transition to an *anti*-conformation at the insertion site (n), allowing it to base pair with an incoming dCTP. According to our NMR results (Figures 3−4; Table 1), the planar FAF at n − 1 exhibited a greater percentage of the *syn*-S-type conformation, while the nonplanar FABP favored the *anti*-B-type conformation (Table 1). It is plausible that the B-conformer of FABP stalls at n − 1, thus resulting in the accumulation of an extension product at the prelesion site. We have reported previously that FAF in the <u>T</u>G*A sequence adopted a higher percentage of the B-conformer and exhibited a stronger stalling effect at the n − 1 site, while the S-conformer of <u>C</u>G*A stalled primer extension at the lesion site (n).[35] In addition, the B-conformer in the FAF-modified <u>T</u>G*A sequence context favored a misinsertion of dATP by 80-fold compared to the <u>C</u>G*A sequence due to the presence of 5′-T.

Lesion and Sequence Effects on Insertion Efficiencies during TLS in the Presence of Kf-exo⁻. Both FABP and FAF lesions exhibited different relative insertion efficiencies (f_{ins}), depending on the primer positions (Table 2). The impact of the lesion was generally minimal at the prelesion (n − 1) and lesion (n) sites. The conformational characteristics of FABP and FAF and their f_{ins} values for the matched (dG*:dC) series were comparable for the G*C<u>A</u> and G*C<u>T</u> sequences (Figure 6). Although both FABP and FAF at the lesion site (n) favored insertion of dCTP, the preference was more prominent for FABP than for FAF. This result may be due to the high S-conformer percentage of the FAF adduct (Table 1). Similar trends in extension efficiency patterns were observed for the dA-mismatched series, although extremely low f_{ins} values were observed for FABP at the lesion site (ND, Table 2). The most dramatic lesion effects were observed at the n + 1 site. We noted a slightly greater lesion effect of FAF over FABP in both G*C<u>A</u> and

G*C<u>T</u> sequence contexts. These results are consistent with the blocking patterns observed in the running start and SPR binding results discussed above and could be attributed to the relatively higher S-conformer population of the FAF adducts compared to the FABP adducts (Table 1).

Figure 9 shows difference factor (DF) histograms to compare the relative insertion efficiency between two sequences (G*C<u>A</u>

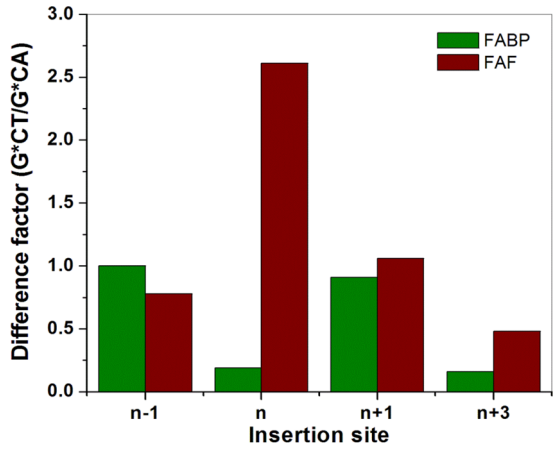


Figure 9. Histogram of the relative insertion efficiency of correct nucleotides between G*CT and G*CA sequences for FABP and FAF adducts. A difference factor ($f_{ins\ G*CT}/f_{ins\ G*CA}$) < 1 signifies a better insertion efficiency in the G*CA sequence compared to that of the G*CT at a particular insertion site, whereas a factor above one has opposite efficiencies.

vs G*C<u>T</u>) at various insertion points in the matched (dG*:dC) series. DF is defined as the f_{ins} of G*C<u>T</u> over G*C<u>A</u> ($f_{ins\ G*CT}/f_{ins\ G*CA}$), and thus, DF below one signifies a better insertion efficiency in the G*C<u>A</u> sequence compared to G*C<u>T</u> at a particular insertion site, while DF above one signifies opposite efficiencies. It is clear that the sequence effect was most dramatic at the lesion site (n), i.e., DF = 0.18 and 2.61 for the FABP and FAF adducts, respectively. We observed that the sequence effects were negligible (DF ≈ 1.0) at the prelesion site (n − 1) (Figure 9 and Supporting Information, Table S9). These results seem to be in line with the NMR results showing that the nature of the 3′-next flanking base (G*C<u>A</u> vs G*C<u>T</u>) did not affect the conformational patterns at n − 1. In the FABP adducts at the lesion site, dATP insertion did not lead to any significant sequence effects (Table 2). In contrast, in the FAF adducts, the G*C<u>T</u> sequence favored dATP insertion by 11-fold compared to that of G*C<u>A</u>. This result is probably due to the relatively higher population of the S-conformer (76%) in the G*C<u>T</u> sequence compared to that of the G*C<u>A</u> (42%). Extension of the dG insertion at the n + 1 site was significantly reduced due to perturbations at the active site of the polymerase. As mentioned above, an extremely low insertion rate was noted for the FABP adducts in the dA-mismatched series. The insertion of dATP at n + 1 was observed for the FAF adducts, but with low efficiency. However, no significant sequence effects were observed at the n + 1 position in either the matched (Figure 9, DF = 0.95 and 1.06 for the FABP and FAF adducts, respectively) or the mismatched series. The FABP in the G*C<u>A</u> series at n + 3 adopted exclusively the B-conformation, yet exhibited a 6-fold greater insertion extension efficiency of dATP than G*C<u>T</u> for the matched series (Figure 9, DF = 0.16). In contrast, the planar FAF showed only a 16% difference in the B:S population ratios (Table 1), thus resulting in a minimal sequence effect (DF = 0.45) (Figure 8).

In summary, we conducted systematic ^{19}F NMR, SPR, and primer elongation kinetic studies for the TLS of the G*C<u>T</u> and G*C<u>A</u> sequences containing the bulky arylamine FAF and FABP adducts. The results are summarized in Figure 10. DNA adducts in the G*C<u>T</u> duplex exhibited greater populations of the S-conformer compared to that in the the G*C<u>A</u> duplex with the same adducts. We also found the 3′-next flanking *T* in the former promotes lesion stacking, thus supporting the greater S conformation. These lesions exhibited unique sequence-dependent conformational heterogeneities at various elongation positions, including the replication fork that differentially contributes to template−primer bindings and insertion efficiency during TLS. In addition, the SPR binding results revealed that the presence of adduct in the stacked conformation decreases the binding affinity of the complementary strands in the order of unmodified > FABP > FAF and G*C<u>A</u> > G*C<u>T</u>. Full-length primer extensions across the modified templates were significantly faster in the G*C<u>A</u> sequence, which commonly adopted the B-conformation. Primer extension was stalled at (n) and near (n − 1 and n + 1) the lesion site, and the extent of blockage and the extension rates across the lesion were influenced by the DNA sequences as well as the nature of the lesion (FAF vs FABP) and the polymerase employed (Kf-exo- vs Dpo4). For example, with Kf-exo$^-$ the *anti*-B-conformeric FABP showed stalling at both the prelesion (n − 1) and lesion (n) sites, whereas the *syn*-S-conformeric FAF showed stalling at all three sites (n − 1, n, and n + 1). Moreover, the relative nucleotide insertion rates (f_{ins}) were small at n − 1 but were significantly reduced immediately 5′ upstream of the lesion site (n + 1 and n + 3). Together, these results demonstrated the roles of lesion-induced conformational heterogeneity in modulating the efficiency of TLS.

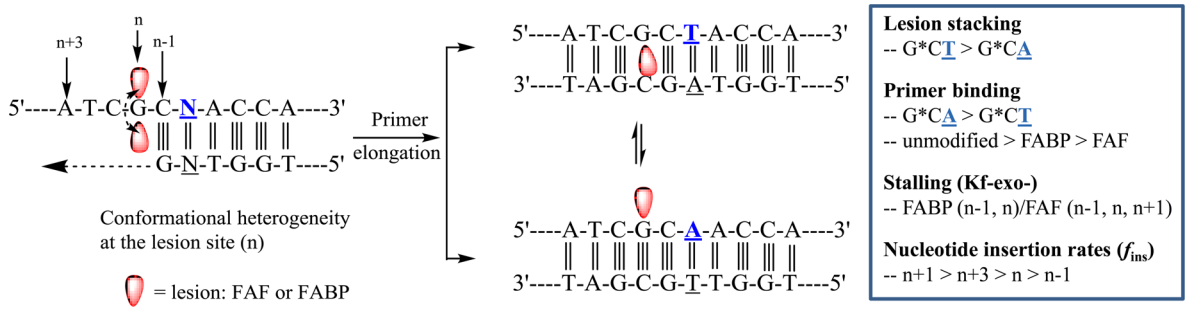


Figure 10. Summary of present results in terms of lesion stacking, primer binding, lesion stalling, and nucleotide insertion rates.

dx.doi.org/10.1021/bi5003212 | *Biochemistry* 2014, 53, 4059−4071

■ ASSOCIATED CONTENT

Supporting Information

Imino proton and full dynamic ^{19}F NMR spectra of FABP-modified 11-mer duplexes, MALDI-TOF spectra of biotinylated templates, and simulated SPR sensograms; kinetic parameters for dNTP insertion at different primer positions; and the dissociation constant values (k_d) obtained from SPR binding experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: 401-874-5024. Fax: 401-874-5766. E-mail: bcho@uri.edu.

Present Addresses

†(V.G.V.) Chemical Laboratory, CSIR-CLRI, Adyar, Chennai 600020, India.

‡(S.P.) CSIR-Indian Institute of Toxicology & Research (IITR), M.G. Marg, Lucknow 226001, India.

Funding

This research is supported by NCI/NIH (CA098296) and NCRR/NIGMS (P20 GM103430-12) (to B.P.C.). V.G.V acknowledges DST, India for the Ramanujan Fellowship.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Drs. Peter F. Guengerich of Vanderbilt University and Catherine Joyce/Olga Potapova of Yale University for providing Dpo4 and Kfexo$^-$ (D424A) for nucleotide insertion kinetic assays.

■ ABBREVIATIONS

FAF, *N*-(2′-deoxyguanosin-8-yl)-7-fluoro-2-aminofluorene; FABP, *N*-(2′-deoxyguanosin-8-yl)-4′-fluoro-4-aminobiphenyl; Kf-exo$^-$, exonuclease deficient Klenow fragment; SPR, surface plasmon resonance; TLS, translesion synthesis

■ REFERENCES

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

(2) Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., and Ellenberger, T., Eds. (2006) *DNA Repair and Mutagenesis*, 2nd ed. ed., ASM Press, Washington, DC.

(3) Hubscher, U., and Maga, G. (2011) DNA replication and repair bypass machines. *Curr. Opin. Chem. Biol. 15*, 627−635.

(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) Federley, R. G., and Romano, L. J. (2010) DNA polymerase: structural homology, conformational dynamics, and the effects of carcinogenic DNA adducts. *J. Nucleic Acids*, 2010.

(6) Beland, F. A., and Kadlubar, F. F. (1990) *Handbook of Experimental Pharmacology*, Spring-Verlag, Heidelberg, Germany.

(7) 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*, 73−114.

(8) Shibutani, S., Suzuki, N., and Grollman, A. P. (1998) Mutagenic specificity of (acetylamino)fluorene-derived DNA adducts in mammalian cells. *Biochemistry 37*, 12034−12041.

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

(10) Cho, B. P., Beland, F. A., and Marques, M. M. (1992) NMR structural studies of a 15-mer DNA sequence from a ras protooncogene, modified at the first base of codon 61 with the carcinogen 4-aminobiphenyl. *Biochemistry 31*, 9587−9602.

(11) 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 19F NMR, UV, and CD spectroscopy. *Biochemistry 46*, 13310−13321.

(12) 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.

(13) O'Handley, S. F., Sanford, D. G., Xu, R., Lester, C. C., Hingerty, B. E., Broyde, S., and Krugh, T. R. (1993) Structural characterization of an N-acetyl-2-aminofluorene (AAF) modified DNA oligomer by NMR, energy minimization, and molecular dynamics. *Biochemistry 32*, 2481−2497.

(14) Zhou, L., Rajabzadeh, M., Traficante, D. D., and Cho, B. P. (1997) Conformational heterogeneity of arylamine-modified DNA: 19F NMR evidence. *J. Am. Chem. Soc. 119*, 5384−5389.

(15) Mao, B., Vyas, R. R., Hingerty, B. E., Broyde, S., Basu, A. K., and Patel, D. J. (1996) Solution conformation of the N-(deoxyguanosin-8-yl)-1-aminopyrene ([AP]dG) adduct opposite dC in a DNA duplex. *Biochemistry 35*, 12659−12670.

(16) 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.

(17) 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.

(18) Cho, B. P., Beland, F. A., and Marques, M. M. (1994) NMR structural studies of a 15-mer DNA duplex from a ras protooncogene modified with the carcinogen 2-aminofluorene: conformational heterogeneity. *Biochemistry 33*, 1373−1384.

(19) Shapiro, R., Ellis, S., Hingerty, B. E., and Broyde, S. (1998) Effect of ring size on conformations of aromatic amine-DNA adducts: the aniline-C8 guanine adduct resides in the B-DNA major groove. *Chem. Res. Toxicol. 11*, 335−341.

(20) Meneni, S. R., D'Mello, R., Norigian, G., Baker, G., Gao, L., Chiarelli, M. P., and Cho, B. P. (2006) Sequence effects of aminofluorene-modified DNA duplexes: thermodynamic and circular dichroism properties. *Nucleic Acids Res. 34*, 755−763.

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

(22) 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.

(23) Geacintov, N. E., Cosman, M., Hingerty, B. E., Amin, S., Broyde, S., and Patel, D. J. (1997) NMR solution structures of stereoisometric covalent polycyclic aromatic carcinogen-DNA adduct: principles, patterns, and diversity. *Chem. Res. Toxicol. 10*, 111−146.

(24) Liu, Y., Reeves, D., Kropachev, K., Cai, Y., Ding, S., Kolbanovskiy, M., Kolbanovskiy, A., Bolton, J. L., Broyde, S., Van Houten, B., and Geacintov, N. E. (2011) Probing for DNA damage with beta-hairpins: similarities in incision efficiencies of bulky DNA adducts by prokaryotic and human nucleotide excision repair systems in vitro. *DNA Repair 10*, 684−696.

(25) Mu, H., Kropachev, K., Wang, L., Zhang, L., Kolbanovskiy, A., Kolbanovskiy, M., Geacintov, N. E., and Broyde, S. (2012) Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies. *Nucleic Acids Res. 40*, 9675−9690.

(26) Jain, V., Hilton, B., Patnaik, S., Zou, Y., Chiarelli, M. P., and Cho, B. P. (2012) Conformational and thermodynamic properties modulate the nucleotide excision repair of 2-aminofluorene and 2-acetylamino-

fluorene dG adducts in the NarI sequence. *Nucleic Acids Res. 40*, 3939−3951.
(27) Vaidyanathan, V. G., Liang, F., Beard, W. A., Shock, D. D., Wilson, S. H., and Cho, B. P. (2013) Insights into the conformation of aminofluorene-deoxyguanine adduct in a DNA polymerase active site. *J. Biol. Chem. 288*, 23573−23585.
(28) Burnouf, D., Koehl, P., and Fuchs, R. P. (1989) Single adduct mutagenesis: strong effect of the position of a single acetylaminofluorene adduct within a mutation hot spot. *Proc. Natl. Acad. Sci. U.S.A. 86*, 4147−4151.
(29) Broschard, T. H., Koffel-Schwartz, N., and Fuchs, R. P. (1999) Sequence-dependent modulation of frameshift mutagenesis at NarI-derived mutation hot spots. *J. Mol. Biol. 288*, 191−199.
(30) Jain, V., Hilton, B., Lin, B., Patnaik, S., Liang, F., Darian, E., Zou, Y., Mackerell, A. D., Jr., and Cho, B. P. (2013) Unusual sequence effects on nucleotide excision repair of arylamine lesions: DNA bending/distortion as a primary recognition factor. *Nucleic Acids Res. 41*, 869−880.
(31) Liang, F., and Cho, B. P. (2010) Enthalpy-entropy contribution to carcinogen-induced DNA conformational heterogeneity. *Biochemistry 49*, 259−266.
(32) 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.
(33) 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.
(34) 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.
(35) Vaidyanathan, V. G., and Cho, B. P. (2012) Sequence effects on translesion synthesis of an aminofluorene-DNA adduct: Conformational, thermodynamic, and primer extension kinetic studies. *Biochemistry 51*, 1983−1995.
(36) 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.
(37) Shapiro, R. U. G., Zawadzka, H., Broyde, S., and Hingerty, B. E. (1986) Conformation of d(CpG) modified by the carcinogen 4-aminobiphenyl: a combined expeimental and theoretical analysis. *Biochemistry 25*, 2198−2205.
(38) 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.
(39) Vaidyanathan, V. G., Xu, L., and Cho, B. P. (2012) Binary and ternary binding affinities between exonuclease-deficient Klenow fragment (Kf-exo$^-$) and various arylamine DNA lesions characterized by surface plasmon resonance. *Chem. Res. Toxicol. 25*, 1568−1570.
(40) Wang, M., Mahrenholz, A., and Lee, S. H. (2000) RPA stabilizes the XPA-damaged DNA complex through protein-protein interaction. *Biochemistry 39*, 6433−6439.
(41) You, J. S., Wang, M., and Lee, S. H. (2003) Biochemical analysis of the damage recognition process in nucleotide excision repair. *J. Biol. Chem. 278*, 7476−7485.
(42) Lebbink, J. H., Fish, A., Reumer, A., Natrajan, G., Winterwerp, H. H., and Sixma, T. K. (2010) Magnesium coordination controls the molecular switch function of DNA mismatch repair protein MutS. *J. Biol. Chem. 285*, 13131−13141.
(43) Sedletska, Y., Culard, F., Midoux, P., and Malinge, J. M. (2013) Interaction studies of muts and mutl with DNA containing the major cisplatin lesion and its mismatched counterpart under equilibrium and nonequilibrium conditions. *Biopolymers 99*, 636−647.
(44) Persson, B., Stenhag, K., Nilsson, P., Larsson, A., Uhlen, M., and Nygren, P. (1997) Analysis of oligonucleotide probe affinities using surface plasmon resonance: a means for mutational scanning. *Anal. Biochem. 246*, 34−44.
(45) Sakao, Y., Nakamura, F., Ueno, N., and Hara, M. (2005) Hybridization of oligonucleotide by using DNA self-assembled monolayer. *Colloids Surf., B 40*, 149−152.
(46) Bates, P. J., Dosanjh, H. S., Kumar, S., Jenkins, T. C., Laughton, C. A., and Neidle, S. (1995) Detection and kinetic studies of triplex formation by oligodeoxynucleotides using real-time biomolecular interaction analysis (BIA). *Nucleic Acids Res. 23*, 3627−3632.
(47) 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.
(48) Ling, H., Sayer, J. M., Plosky, B. S., Yagi, H., Boudsocq, F., Woodgate, R., Jerina, D. M., and Yang, W. (2004) Crystal structure of a benzo[a]pyrene diol epoxide adduct in a ternary complex with a DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A. 101*, 2265−2269.
(49) Pednekar, V., Weerasooriya, S., Jasti, V. P., and Basu, A. K. (2014) Mutagenicity and genotoxicity of (5′S)-8,5′-cyclo-2′-deoxyadenosine in *Escherichia coli* and replication of (5′S)-8,5′-cyclopurine-2′-deoxynucleosides in vitro by DNA polymerase IV, exo-free Klenow fragment, and Dpo4. *Chem. Res. Toxicol. 27*, 200−210.
(50) Gowda, A. S., Krishnegowda, G., Suo, Z., Amin, S., and Spratt, T. E. (2012) Low fidelity bypass of O(2)-(3-pyridyl)-4-oxobutylthymine, the most persistent bulky adduct produced by the tobacco specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone by model DNA polymerases. *Chem. Res. Toxicol. 25*, 1195−1202.
(51) Christov, P. P., Chowdhury, G., Garmendia, C. A., Wang, F., Stover, J. S., Elmquist, C. E., Kozekova, A., Angel, K. C., Turesky, R. J., Stone, M. P., Guengerich, F. P., and Rizzo, C. J. (2010) The C8−2′-deoxyguanosine adduct of 2-amino-3-methylimidazo[1,2-d]-naphthalene, a carbocyclic analogue of the potent mutagen 2-amino-3-methylimidazo[4,5-f]quinoline, is a block to replication in vitro. *Chem. Res. Toxicol. 23*, 1076−1088.
(52) Silverman, A. P., Jiang, Q., Goodman, M. F., and Kool, E. T. (2007) Steric and electrostatic effects in DNA synthesis by the SOS-induced DNA polymerases II and IV of *Escherichia coli*. *Biochemistry 46*, 13874−13881.