Binding of the human nucleotide excision repair proteins XPA and XPC/HR23B to the 5R-thymine glycol lesion and structure of the cis-(5R,6S) thymine glycol epimer in the 5'-GTgG-3' sequence: destabilization of two base pairs at the lesion site

Kyle L. Brown1, Marina Roginskaya2, Yue Zou2, Alvin Altamirano3, Ashis K. Basu3 and Michael P. Stone1,*

1Department of Chemistry and Center in Molecular Toxicology, Vanderbilt University, Nashville, TN 37235, 2Department of Biochemistry and Molecular Biology, James H. Quillen College of Medicine, East Tennessee State University, Johnson City, TN 37614 and 3Department of Chemistry, University of Connecticut, Storrs, CT 06269, USA

Received September 4, 2009; Revised September 21, 2009; Accepted September 22, 2009

ABSTRACT

The 5R thymine glycol (5R-Tg) DNA lesion exists as a mixture of cis-(5R,6S) and trans-(5R,6R) epimers; these modulate base excision repair. We examine the 7:3 cis-(5R,6S):trans-(5R,6R) mixture of epimers paired opposite adenine in the 5'-GTgG-3' sequence with regard to nucleotide excision repair. Human XPA recognizes the lesion comparably to the C8-dG acetylaminofluorene (AAF) adduct, whereas XPC/HR23B recognition of Tg is superior. 5R-Tg is processed by the Escherichia coli UvrA and UvrABC proteins less efficiently than the C8-dG AAF adduct. For the cis-(5R,6S) epimer Tg and A are inserted into the helix, remaining in the Watson–Crick alignment. The Tg N3H imine and A N6 amine protons undergo increased solvent exchange. Stacking between Tg and the 3'-neighbor G+C base pair is disrupted. The solvent accessible surface and T2 relaxation of Tg increases. Molecular dynamics calculations predict that the axial conformation of the Tg CH3 group is favored; propeller twisting of the Tg•A pair and hydrogen bonding between Tg OH6 and the N7 atom of the 3'-neighbor guanine alleviate steric clash with the 5'-neighbor base pair. Tg also destabilizes the 5'-neighbor G+C base pair. This may facilitate flipping both base pairs from the helix, enabling XPC/HR23B recognition prior to recruitment of XPA.

INTRODUCTION

5,6-dihydroxy-5,6-dihydro-2'-thymine, thymine glycol (Tg), is formed by exposure to radiation and chemical oxidants (1,2). It is also formed by oxidation of 5-methylcytosine to 5-methylcytosine glycol, followed by deamination (3,4). The C5 and C6 atoms of Tg are chiral and it exists in DNA as diastereomeric pairs of epimers, the 5R cis, trans pair (5R,6S;5R,6R) and the 5S cis, trans pair (5S,6R; 5S,6S) (Scheme 1) (5–7). The 5R pair is more abundant and more stable; in either case, the cis isomers predominate at the nucleoside level (6). Human cells repair hundreds of Tg lesions per day (8,9). The 5R-Tg lesion has been examined in the 5'-ATgA-3' sequence, paired opposite dA (10). It was concluded that Tg was partially extrahelical (10). It was also reported that the structure of 5R-Tg placed opposite dA in the 5'-GTgC-3' sequence was disordered (11). These studies (10,11) did not report cis–trans epimerization (5–7) of Tg. In DNA the complementary base modulates the cis–trans equilibrium of 5R-Tg. When paired opposite dA in the 5'-GTgG-3' sequence a 7:3 cis (5R,6S); trans (5R,6R) mixture exists at equilibrium at 25°C, whereas
It is a substrate for base excision repair. This is human NER (20).

Here we report on the 7:3 \( \text{cis}-(5R,6S) \) and \( \text{trans}-(5R,6R) \) Tg lesions. When the \( 5R \)-Tg isomer is paired opposite dA in this \( 5' \)-GTgG-3' sequence a 7:3 \( \text{cis}-(5R,6S) \); \( \text{trans}-(5R,6R) \) mixture is present at equilibrium, in slow exchange on the NMR timescale (12).

### Scheme 1.

(A) Interconversion of the \( \text{cis}-(5R,6S) \) and \( \text{trans}-(5R,6R) \) Tg lesions. When the \( 5R \)-Tg isomer is paired opposite dA in this \( 5' \)-GTgG-3' sequence a 7:3 \( \text{cis}-(5R,6S) \); \( \text{trans}-(5R,6R) \) mixture is present at equilibrium, in slow exchange on the NMR timescale (12).

(B) Oligodeoxynucleotide duplex used for NMR studies, indicating the numbering of the nucleotides. X\(^{0} \) is the \( \text{cis}-(5R,6S) \) Tg lesion.

when paired with dG in the same sequence, only the \( \text{cis} \) (5R,6S) epimer is observed (12). When \( \text{cis}-(5R,6S) \) Tg is mismatched with dG in the \( 5' \)-GTgG-3' sequence it assumes the wobble orientation, and shifts toward the major groove. This increases its solvent accessible surface but it remains stacked into the helix. Intrastrand hydrogen bonding between the hydroxyl on C6 of Tg and the N7 of the 3' purine (13) is weak (14).

The \( 5R \)-Tg lesion hinders DNA replication (15,16). It is a substrate for base excision repair. This is mediated by at least two DNA N-glycosylase/AP lyases that are influenced by the diastereoisomer of Tg, the \( \text{cis}-\text{trans} \) epimerization of each diastereoisomer, and the identity of the complementary purine (17). The \( 5R \)-Tg lesion is also repaired by nucleotide excision repair (NER), although the effects of the \( \text{cis}-\text{trans} \) epimerization of each diastereoisomer of Tg with regard to NER have not been characterized. Both randomly introduced \( 5R \)-Tg and abasic sites are substrates for the *Escherichia coli* UvrABC proteins (18,19). Tg is also excised in *vivo* by human NER (20). Here we report on the 7:3 \( \text{cis}-(5R,6S):\text{trans}-(5R,6R) \) mixture of Tg epimers paired opposite adenine in the \( 5' \)-GTgG-3' sequence (12) with regard to NER by the NER proteins of *E. coli* and the binding of the lesion by the human NER proteins XPA and XPC/HR23B. The \( 5R \)-Tg lesion is a good substrate for binding by UvrA and excision by UvrABC, although the bulky C8-dG DNA adduct of AAF is recognized and incised more efficiently. However, recognition of \( 5R \)-Tg by the human XPC/HR23B complex is superior to the bulky AAF adduct, whereas recognition by human XPA is comparable to the AAF adduct. To elucidate structure–activity relationships underlying these observations, the structure of the \( \text{cis}-(5R,6S) \) Tg epimer embedded in the same \( 5' \)-GTgG-3' sequence and placed opposite to deoxyadenosine has also been refined. It remains in the Watson–Crick orientation with respect to the complementary dA, but the solvent accessible area of Tg is increased. The complementary A\(^{19} \) remains stacked in the helix. Significantly, the \( 5R \)-Tg lesion also destabilizes the 5'-neighbor G\( \text{\textasciitilde} \text{C} \) base pair. We propose that this lowers the activation barrier with respect to flipping both base pairs from the helix, at least in this particular sequence, enabling XPC/HR23B to recognize \( 5R \)-Tg prior to the recruitment of XPA.

### METHODS

**Sample preparation**

The undamaged ND-50-bp and AAF-50-bp oligodeoxynucleotides were constructed as described (21). The Tg-modified \( 5'\text{d}(\text{GTGCCTGTTTGT})-3' \) (17) was characterized using a Voyager-DE MALDI-TOF mass spectrometer (PerSeptive Biosystems, Inc., Foster City, CA, USA). It was ligated with a \( 5' \),\(^{32} \text{P} \)-labeled 20-mer, and a 19-mer on the \( 5' \) and 3' -ends, respectively, to form the \( 5' \),\(^{32} \text{P} \)-labeled Tg-51mer. The oligodeoxynucleotides \( 5'\text{d}(\text{GTGCCTGTTTGT})-3' \) and \( 5'\text{d}(\text{ACAAACACGACAC})-3' \) were purchased from the Midland Certified Reagent Co. (Midland, TX, USA) and purified by reverse phase HPLC.

**Binding as measured by electrophoretic mobility shift assays**

Assays for the UvrA protein were performed as described (21). For the human NER proteins XPA and XPC/HR23B, the binding assay was performed as described for XPA binding (22), with minor modifications. The radioactivity of DNA bands on gels was quantified using a Fuji FLA-5000 phosphoimager. Dissociation constants...
(K_d) were determined from the binding curves as a concentration of the substrate at which half of DNA was bound to protein. Three experiments were performed for each binding curve.

**UVRABC incision assays**

The 5'-32P-labeled substrates (2 nM) were incised by UvrABC nucleases (UvrA, 15 nM; UvrB, 250 nM; UvrC, 100 nM) in the UvrABC buffer with 1 mM ATP at 37°C and the products were resolved on a 12% polyacrylamide gel and analyzed as described (21). The initial rates were calculated by linear least-squares fits of data. The substrate incised in femtomoles was calculated based on the total molar amount of substrate employed in each reaction and the incision percentage of the substrate. Three experiments were performed to determine initial rates.

**Nuclear magnetic resonance spectroscopy**

Oligodeoxynucleotides were annealed in 20 mM sodium phosphate, containing 100 mM NaCl, 10 μM NaN_3 and 50 μM Na_2EDTA (pH 7.0). Experiments were performed at 800 MHz. NOE spectra for the nonexchangeable protons were recorded at 30°C with mixing times of 80, 150, 200 and 250 ms using States-TPPI phase cycling, 512 real data points in the d1 dimension with 32 scans per FID, 2K real data points in the d2 dimension, sweep width of 10 p.p.m., and a relaxation delay of 2.0 s. The water resonance was suppressed using presaturation. NOE spectra of exchangeable protons were obtained at 5°C using watergate H_2O suppression (23). T_1 spin–lattice relaxation experiments were collected using the inversion recovery method (24,25). T_2 transverse relaxation experiments were collected using the CPMG method (24,25). 1H spectra were referenced to 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid, sodium salt (3-TMSP). The program XWINNMR (Bruker Inc., Billerica, MA, USA) was used for data processing. A skewed sinesbell-squared apodization function with a 90° phase shift was used for NOE experiments; the same function with a 10° phase shift was used for COSY experiments.

**Distance restraints**

NOE intensities were determined from volume integration using SPARKY (26). Intensities for the cis-(5R, 6S) epimer were corrected for the 7:3 cis-(5R,6S):trans-(5R,6R) molar ratio. These were combined with intensities generated from CORMA analysis of a B-form starting structure producing a hybrid intensity matrix (27,28), which was refined using MARDIGRAS (29–31) with the RANDMARDI function. Calculations at mixing times of 80, 150, 200 and 250 ms were run at correlation times of 2, 3, 4 and 5 ns. Distance restraints were divided into categories indicative of confidence levels. Empirical restraints were used to define Watson–Crick hydrogen bonding, but not for the Tg 6A 19 or G 5 7C 20 pairs.

**Torsion angle restraints**

3J 1H coupling constants were obtained by amplitude constrained multiplet evaluation of COSY data (32). Electronegativity of substituent Karplus curves were generated and converted to phase angle space assuming a maximum pucker amplitude (Φ) of 44° (33,34). Scalar couplings were fit to the curves to determine pseudo-rotation ranges (P). The sugar pseudorotation and amplitude ranges were converted to restraints for the dihedral angles ω_0 to ω_4. Measurements of the mol fraction of sugar puckerers in the N or S conformations were determined from the sum of J_122' and J_122' scalar couplings (34). Nucleotides with <50% X_5 were allowed to explore both N and S conformations during rMD calculations (ρ = 0°–210°). Nucleotides with X_5 > 50% were restrained such that ρ = 125°–210°. Backbone torsion angles were restrained with data where available; otherwise they were restrained empirically based on canonical A-form and B-form values.

**Structural refinement**

Partial charges for Tg were generated using GAUSSIAN (35). Geometry optimization and ESP calculations were performed using the Hartree–Fock method with the 6-31G* basis set (12). The output was formatted using ANTECHAMBER (36). Starting structures were generated using NAB (37) and energy minimized using the SANDER module of AMBER (36). Coordinate and topology files were generated with xLeap (36) using ff99 parameters (38). The restraint function utilized square-well potentials (39). The generalized Born model was used for simulated annealing calculations, with a salt concentration of 0.2 mM (40,41). Temperature was maintained using the Berendsen algorithm (42). Complete relaxation matrix analysis (27,28) was performed to determine agreement with 1H NOE data. A refined structure from simulated annealing was neutralized with the addition of sodium ions and placed in a truncated octahedral TIP3P water box with periodic boundaries at a distance of 8 A from the solute. After equilibration, a 10-ns isothermal rMD calculation was performed. The temperature was controlled using the Langevin thermostat (43,44) with a collision frequency of 1 ps^{-1}. Electrostatic interactions were treated with the PME method (45). A 15-A cutoff for nonbonded interactions was used. Bond lengths involving hydrogen were fixed using SHAKE (46). A structural ensemble was extracted from the isothermal trajectory using PTRAJ (36). The heavy atoms were subjected to pairwise RMSD comparisons using SUPPOSE. Helicoidal analysis was performed using CURVES (47,48). Structures were rendered using Chimera (49). Solvent-excluded and solvent-accessible surface areas of individual bases as a function of probe radius were calculated using MSMS (50).

**RESULTS**

**Dodecamer containing the 5R-Tg lesion**

The 5'-d(GTGCGXGTTTGGT)-3', X = 5R-Tg, was subject to mass spectrometric analysis, which yielded the anticipated molecular ion peak with mass 3732 (m/z). Capillary gel electrophoretic and HPLC analysis showed that it eluted as a single peak. The adducted
oligodeoxynucleotide was pure and existed as a single chromatographically separable species.

Recognition and incision of the 5R-Tg substrate by E. coli UvrABC proteins

A 51-bp substrate containing 5R-Tg (Supplementary Scheme S1) was utilized. A 50-bp substrate containing an AAF adduct (AAF-50 bp) that is recognized and incised by E. coli UvrABC (21), was used for comparison. A 50-bp substrate (ND-50 bp) was used as a negative control. Figure 1A and B show gel mobility shift assays for UvrA binding to Tg and AAF substrates, respectively. The slowly migrating bands represent the formation of DNA–UvrA2 complexes. Binding isotherms (Figure 1C) showed that the affinity of UvrA to the 5R-Tg substrate was high, though the binding of UvrA to the AAF substrate was stronger (Table 1). The UvrA had a 2.4-fold lower affinity to the Tg-50-bp substrate ($K_d = 24 \pm 2 \text{ nM}$) than the AAF-51-bp substrate ($K_d = 10 \pm 1 \text{ nM}$). The $K_d$ for UvrA binding to the AAF substrate agreed with the reported value (21). UvrA had lower affinity to the undamaged than to the damaged substrate, in agreement with previous observations (21). While the differing sequence contexts of the Tg- and AAF–DNA substrates (Scheme 2A) might affect their interactions with the UvrABC system (51–53), it was assumed that sequence specific effects were smaller than those caused by the difference in the types of DNA lesions.

UvrC cut the damaged strand 4 or 5 nt 3' to the lesion and then 8 nt 5' to the lesion (Supplementary Scheme S1). The substrates were 32P-labeled at the 5'-end and the incision products were 18-mers. Figure 2A and B show temporal data for UvrABC incision of Tg and AAF substrates, respectively. No products were observed for the ND-50-bp undamaged substrate. The initial rates were determined from the relative yields of the products (Figure 2C). For the 5R-Tg substrate the initial rate was 0.48 ± 0.04 fmol/min. For the AAF substrate the comparable rate was 0.80 ± 0.02 fmol/min. Thus, the 5R-Tg substrate was incised 1.7X less efficiently than was the AAF substrate. However, the incision rate of the 5R-Tg substrate was greater than that of the DNA helix-distorting cross-linked tandem G(8,5-Me)T lesion (21). Therefore, in this sequence 5R-Tg was a good substrate for the E. coli UvrABC proteins.

| Protein        | $K_d$ for TG-51 bp (nM) | $K_d$ for AAF-50 bp (nM) |
|----------------|-------------------------|--------------------------|
| UvrA           | 24 ± 2                  | 10 ± 1                   |
| XPA           | 48 ± 4                  | 44 ± 6                   |
| XPC/HR23B     | 18 ± 2                  | 27 ± 3                   |

aData represent the means ± SD of three experiments.

Table 1. Equilibrium dissociation constants for binding of UvrA and human NER proteins to TG-51-bp and AAF-50-bp DNA substrates at 30°C

Binding of human NER proteins to the 5R-Tg substrate

The binding of XPA and XPC/HR23B to the 5R-Tg damaged duplex was compared with the AAF-damaged duplex using EMSA (Figure 3). Neither XPA nor XPC/HR23B bound to the undamaged substrate. The shifted bands represent formation of DNA–XPA2 complexes. At low concentrations XPA bound to AAF-50 bp with a greater affinity than to Tg-51 bp, which was evident from a comparison of the 25 nM XPA lanes for 5R-Tg and AAF. At concentrations >50 nM, XPA bound to both substrates with comparable affinities (Figure 3A). The binding of XPC/HR23B to 5R-Tg and AAF-damaged...
DNA is shown in Figure 3B. Figure 3C shows binding isotherms for XPA interacting with the 5R-Tg- and AAF–DNA substrates. The dissociation constants estimated for the 5R-Tg and AAF substrates were similar, 48 ± 4 nM and 44 ± 6 nM, respectively (Table 1). XPC/HR23B bound to Tg-51 bp with a greater affinity than to AAF-50 bp (Figure 3D). Unlike XPC/HR23B the E. coli NER protein UvrA bound better to AAF–DNA adduct than to Tg-DNA adduct (Figure 1).

Nuclear magnetic resonance spectroscopy

Data were collected upon preparation of the duplex containing the 5R-Tg lesion (Scheme 1), and again after 4 weeks. Spectral changes were not observed, suggesting that the cis-(5R,6S) and trans-(5R,6R) epimers had achieved equilibrium. The cis-(5R,6S) epimer was favored 7:3 over the trans-(5R,6R) epimer, as determined by integration of Tg CH₃ peaks (12). It was possible to obtain spectroscopic data for the cis-(5R,6S) epimer; the trans-(5R,6R) epimer was not present at sufficient levels to allow evaluation of its spectrum.

Nonexchangeable DNA protons. Figure S1 in the Supplementary Data shows NOESY cross-peaks between the base aromatic H6/H8 protons and the deoxyribose H1' protons. The resonances were assigned using standard strategies (54,55). Complete sequential NOE connectivity was obtained for both the modified and the complementary strands. With the exception of several of the H4' protons, and the stereotopic assignments of the H5' and H5'' sugar protons, assignments of the deoxyribose protons were made unequivocally; the resonance assignments have been reported (12).

Exchangeable DNA protons. Supplementary Figure S2 shows NOESY cross-peaks between the purine N1H and pyrimidine N3H imino protons and the amino protons of the complementary bases, involved in Watson–Crick hydrogen bonding. At the mismatched X6•A19 pair, the Tg N3H imino resonance was not identified. This was attributed to rapid exchange with solvent. The assignments of the remaining hydrogen-bonded imino and amino protons were made using standard methods (56). The G5 N1H imino resonance was broad at 5°C and disappeared when the temperature was increased to 15°C. In contrast, for an unmodified duplex, the G5 N1H imino resonance was sharp and was observed at temperatures as high as 40°C (12). There was no cross peak between the G5 N1H resonance and G11 N1H, located at base pair C6•G21. This was attributed to its exchange with solvent. The imino resonances for base pairs T6•A23, G5•C22, C6•G21, G7•C18, T8•A17, T9•A16, T10•A15, and G11•C14 were observed. The imino resonances for the terminal base pairs G1•C24 and T12•A13 were not observed, attributed to exchange broadening with water. The assignments have been reported (12).

Tg protons. Figure 4 shows NOESY data obtained for the Tg CH₃ and Tg H6 protons. The proximate Tg H6 and Tg CH₃ protons yielded a strong Tg H6→Tg CH₃ NOE at all mixing times. The G5 H1'→X6 H6 and G5 H8→X6 H6 NOEs were diagnostic of the cis-(5R,6S) configuration. The cis-(5R,6S) Tg CH₃ protons exhibited a chemical shift of 0.49 p.p.m., while the Tg H6 proton resonated at 4.58 p.p.m. A total of 23 NOE cross peaks were assigned between Tg CH₃ and H6 in the cis-(5R,6S) epimer and DNA (seven for Tg H6 and sixteen for Tg CH₃). The trans-(5R,6R) Tg CH₃ protons exhibited a chemical shift of 1.24 p.p.m., while the Tg H6 proton resonated at 4.91 p.p.m. For the trans-(5R,6R) epimer, only one NOE cross peak was observed, between Tg H6 and Tg H2'. Integration of the respective Tg H6→Tg H2' cross peaks confirmed the 7:3 ratio of epimers. A single set of resonances was observed for G5 and G7. Thus, the equilibrium mixture of epimers did not influence the chemical shift environment of the neighboring nucleotides. The spectroscopic assignments have been reported (12).

Spin–lattice relaxation

The T₁ relaxation times of the T₂, T₄, T₉, T₁₀, & T₁₂ CH₃ and X₆ CH₃ groups were compared to those of the unmodified duplex (Supplementary Figure S3). The X₆
CH₃ T₁ relaxation time was 1.9 s faster than was the corresponding T₈ CH₃ relaxation time of the unmodified duplex containing T₈•A⁹. The X₆ CH₃ relaxed an average of 1.4 s faster than other thymine CH₃ groups in the X₆•C₁⁵A₁⁹ duplex.

Spin–spin relaxation
The T₂ relaxation times of the T², T⁸, T⁹, T¹⁰, & T¹² CH₃ and X₆ CH₃ groups were compared to those of the unmodified duplex (Supplementary Figure S4). The Tg CH₃ of the X₆•A¹⁹ duplex relaxed ~70 ms faster than the corresponding CH₃ of the unmodified sample.

Structural refinement
Thirty starting structures were generated, of which half had the Tg CH₃ group in the axial conformation and half had the Tg CH₃ group in the equatorial conformation. These exhibited a maximum pairwise RMSD of 4.3 Å. The distance and torsion angle restraints used in the calculations are summarized in Supplementary Table S1. As Watson–Crick hydrogen bonding at the X₆•A¹⁹ base pair was not observed, hydrogen bond restraints between X₆ and A¹⁹ were not used. All but three of the 30 structures emerged from simulated annealing calculations with Tg CH₃ in the axial conformation. A refined structure with Tg CH₃ in the axial conformation was placed into a truncated octahedron TIP3P water box, and subjected to 10 ns of isothermal rMD calculations at 300 K (Supplementary Figure S5). At 4.38 ns the Tg CH₃–C₅–C₆–H₆ torsion angle shifted from −40° to 50°. This corresponded to a change from the Tg CH₃ axial to the equatorial conformation. The glycosyl Tg O⁴–H¹′–N¹–C₂ torsion angle fluctuated between −90° and −140°.

Figure 3. Binding of human XPC/HR23B and XPA proteins to the site-specific 5R-Tg paired with dA substrate. In (A) and (B), XPA (A) or XPC/HR23B (B) proteins at the indicated concentrations were incubated with 4-nM substrates at 30°C for 30 min in XPA binding buffer and then analyzed on 3.5% polyacrylamide native gels by gel mobility shift assays (XPA, 4°C; XPC/HR23B, room temperature). (C) and (D) show the binding curves generated from the data in (A) and (B).
during the first 4.38 ns; when the Tg CH₃ group shifted from axial to equatorial conformation, this angle fluctuated between -130° and -150°.

The distances between Tg HO5-G7 N7 (Supplementary Figure S5, panel C) and Tg HO6-G7 N7 (Supplementary Figure S5, panel D) were monitored during the isothermal calculations. Analysis of trajectories indicated intra-strand hydrogen bond stabilization of the X₈•A¹⁹ base pair as predicted (13). Hydrogen bond occupancy was defined as hydrogen bond donors and acceptors being within 3.5 Å with an angle cutoff of 120°. During the first 4.38 ns, in which the Tg CH₃ group was in the axial conformation, the Tg HO6-G7 N7 hydrogen bond criteria were satisfied (Supplementary Figure S5, panel D). This induced altered propeller twist at the lesion site. There was 42% occupancy of the Tg OH6-G7 N7 hydrogen bond during the trajectory and the entirety of this occupancy occurred during the first 4.38 ns. In contrast, during the first 4.38 ns, in which the Tg CH₃ group was in the axial conformation, the distance between Tg HO5 and G7 N7 was 4.5 Å, which was not consistent with hydrogen bonding. As a consequence of Tg CH₃ reorienting to the equatorial conformation at 4.38 ns, Tg HO5 shifted to the axial conformation. This allowed for hydrogen bond formation with G7 N7. The Tg OH5-G7 N7 occupancy was 11% over 10 ns and was only observed after the conformational shift. During these isothermal calculations, the RMSD of the DNA backbone heavy atoms (Supplementary Figure S5, panel F) was 2 Å.

Structural ensembles representing both axial and equatorial conformations of Tg CH₃ were extracted for analyses (Supplementary Figure S6). The axial ensemble (PDB ID 2KH5) was extracted from the 120 ps before the conformation change at 4.38 ns; this ensemble had an RMSD of 0.56 Å for the core 6 bp. The equatorial ensemble (PDB ID 2KH6) was extracted 1 ns after the conformational change at 4.38 ns. This had an RMSD of 0.69 Å for the core 6 bp.

To evaluate the accuracies of the rMD ensembles with respect to the ¹H NOESY data, complete relaxation matrix analyses (27,28) were performed (Supplementary Figure S7). The ensemble with Tg CH₃ in the axial conformation had an average sixth root residual Rₓ value of 8.12×10⁻². The Tg lesion exhibited an Rₓ value of 7.19×10⁻² for inter-nucleotide cross-peaks with G³ and 7.61×10⁻² for intra-nucleotide cross-peaks. The ensemble with Tg CH₃ in the equatorial position had an average Rₓ value of 8.15×10⁻². The Tg lesion exhibited an Rₓ value of 1.77×10⁻¹ for inter-nucleotide cross peaks with G³ and 1.01×10⁻¹ for intra-nucleotide cross-peaks. This indicated that both ensembles of structures satisfied the ¹H NOESY data, although at the lesion site, the agreement between the structures with the Tg CH₃ axial conformation and the NOESY data was improved.

**Refined structures**

In both the Tg-axial and Tg-equatorial ensembles the cis-(5R,6S) lesion was positioned in the Watson–Crick orientation such that Tg NH₃ was proximate to A¹⁹ N1 and Tg O⁴ was proximate to A¹⁹ N⁶H (Figure 5). However, for the Tg-axial ensemble, the base pair opening at X₈•A¹⁹ increased such that the Tg shifted toward the major groove. For the Tg-equatorial ensemble, Tg maintained Watson–Crick positioning. Irrespective of the conformation of Tg, the complementary dA remained in the anti conformation about the glycosyl bond and stacked into the duplex (Figure 6). The Tg base stacked below the 5’-neighbor base G⁵, while its complement A¹⁹ stacked below the 5’-neighbor base C²⁰. In the 3'-direction, base stacking between Tg and the G⁷•C¹⁸ base pair was disrupted (Figure 5). Helicoidal analyses of both the axial and equatorial conformations of Tg indicated disturbances in shear, opening, stagger and stretch at the lesion (Figure 6). For the Tg CH₃ axial conformation the disturbance of the glycosyl torsion angle (γ) alleviated steric interactions between Tg CH₃ and the 5’-neighbor purine G³. The solvent accessible surface of Tg in the duplex relative to that of a free Tg base was compared with bases A²³, T⁹, and T¹² (Supplementary Figure S8). Irrespective of whether the Tg CH₃ group was in the axial or equatorial conformation, Tg exhibited an increased solvent accessible surface relative to T⁹ and A²³ but was
Figure 5. Base pair stacking interactions of the cis-(5R,6S) Tg lesion. Comparison stacking interactions in which Tg CH₃ is in the axial (PDB ID 2KH5) or equatorial (PDB ID 2KH6) conformations. The top panel shows the G⁵/C²⁰ base pair (black) stacked above the X⁶/C¹⁵ A¹⁹ base pair (Tg is colored red and A¹⁹ is colored gray). The center panel shows the orientation of the X⁶ lesion (red) with respect the complementary nucleotide A¹⁹ (black). The bottom panel shows the X⁶/C¹⁵ A¹⁹ base pair (Tg is colored red and A¹⁹ is colored black) stacked above the G⁷/C₁⁸ base pair (grey).

Figure 6. The cis-(5R,6S) Tg lesion at the X⁶/A¹⁹ base pair as viewed from the major groove showing potential hydrogen bonding interactions as predicted from analyses of rMD trajectories. (A) The Tg OH₆ formed a hydrogen bond with G⁷ N7 when Tg CH₃ was in the axial conformation (PDB ID 2KH5). (B) When Tg CH₃ was in the equatorial conformation Tg OH₆ did not hydrogen bond with G⁷ N7, however, improved hydrogen bonding was observed with Tg OH₅ (PDB ID 2KH6).
not as exposed to solvent as the terminal T\textsuperscript{12} base. The average accessible surface area of the Tg base was 19\% in the CH\textsubscript{3} axial structure and 18\% in the CH\textsubscript{3} equatorial structure.

DISCUSSION

Defects in human NER are associated with the disease xeroderma pigmentosum (57,58). In human global genome NER, the XPC/HR23B complex (59–61) is the damage recognition factor. Damage-induced thermal destabilization of the helix is believed to modulate recognition of a diverse group of damages by XPC (62–66). The XPA protein is essential for NER and Yang et al. (67) reported that it exists as a homodimer either in the free state or as a complex with human RPA. It binds to mismatched DNA bubble substrates with or without DNA adducts, including the C8-dG adducts of AF, AAF and 1-nitropyrene, and the T[6,4]T photoproduts (68). XPA is likely involved in the verification of DNA damage (64,69). It also probably plays a role in repair factors and stabilization of repair intermediate structures since it binds more efficiently to undamaged ds–ssDNA junctions with ssDNA branches (68), intermediate structures found in NER. In light of observations that the repair of 5'R-Tg by DNA N-glycosylases/AP lyases is modulated by cis–trans epimerization (17) it was of interest to examine the 7:3 cis-(5'R,6'S): trans-(5'R,6'R) mixture of Tg epimers paired opposite adenine in the 5'-GTgG-3' sequence (12) with regard to NER by the NER proteins of E. coli (18,19), and the binding of the lesion by the human NER proteins XPA and XPC/HR23B (20), and to compare these observations with the structure of the cis-(5'R,6'S) Tg lesion in this same sequence.

When placed opposite dA in the 5'-GTgG-3' sequence the 5'R-Tg lesion exists at 25\°C as a 7:3 cis-(5'R,6'S): trans-(5'R,6'R) mixture at equilibrium; the structure of the cis-(5'R,6'S) epimer has been refined herein. That there is no disruption of sequential NOE connectivity for either the complementary strand or the modified strand (Supplementary Figure S2) supports the conclusion that the cis-(5'R,6'S) Tg lesion minimally distorts the helical backbone (12). At the X\textsuperscript{6}•A\textsuperscript{19} pair, both Tg and A\textsuperscript{19} are inserted into the helix (Figure 6). The X\textsuperscript{6}•A\textsuperscript{19} pair remains in a Watson–Crick type alignment in which Tg O\textsuperscript{4} is proximate to the exocyclic amine of A\textsuperscript{19}, and Tg N3H is proximate to A\textsuperscript{19} N1 (Figure 5). Both the exchangeable Tg amine and A\textsuperscript{19} N\textsuperscript{6} amine resonances undergo increased exchange with solvent. In the 3'-direction, stacking between the cis-(5'R,6'S) Tg\textsuperscript{6} and the G\textsuperscript{7}•C\textsuperscript{18} base pair is also disrupted (Figure 6). For the X\textsuperscript{6}•A\textsuperscript{19} pair the cis (5'R,6'S) Tg\textsuperscript{6} lesion is more exposed to solvent, as compared to a Watson–Crick T•A base pair (Supplementary Figure S8). This agrees with findings in the 5'-AXA-3' sequence when the 5'R Tg lesion was placed opposite dA (10). Tg\textsuperscript{6} lesion remains stacked into the duplex and it is not flipped into the major groove (Figure 6).

Nuclear magnetic resonance (NMR) does not differentiate between axial versus equatorial conformations of the Tg CH\textsubscript{3} group. Both are observed in the rMD calculations. The axial conformation provides a modestly improved agreement with the NOE restraints. This is consistent with the crystal structure at the nucleoside level (70). Moreover, quantum mechanical calculations of the modified base predict the axial conformation of the cis-(5'R,6'S) Tg epimer is favored (13). It was also predicted that the Tg•A pair should be stabilized by an intra-strand hydrogen bond between the Tg OH\textsubscript{6} and the N7 position of a 3' purine (13). Its presence is supported by analyses of the present rMD trajectories. This hydrogen bond is present when Tg CH\textsubscript{3} is in the axial conformation; its occupancy in the 10-ns trajectory is 46\% (Supplementary Figure S5, panel D). It induces propeller twist at the lesion site, which alleviates steric interactions between the CH\textsubscript{3} group and the 5'-neighbor guanine. This is reflected in the perturbation of the glycosyl torsion angle \( \chi \) at Tg (Supplementary Figure S5). The rMD trajectories suggest the potential for Tg OH\textsubscript{5}→G\textsuperscript{7} N7 hydrogen bond formation when the Tg CH\textsubscript{3} group shifts from the axial to the equatorial conformation. This alleviates steric clash between the Tg CH\textsubscript{3} group and the 5'-neighbor guanine, and cannot be ruled out. However, the rMD trajectories predict only 11\% occupancy of this hydrogen bond. Overall, it seems that the axial conformation of the cis-(5'R,6'S) Tg epimer is favored. The increase in longitudinal relaxation for the Tg CH\textsubscript{3} protons is consistent with this conclusion; in the axial conformation the Tg CH\textsubscript{3} group faces into the major groove with fewer sources of longitudinal relaxation. Epimerization to the trans-(5'R,6'R) Tg lesion, for which the equatorial conformation of the Tg CH\textsubscript{3} group is favored by 4 kcal/mol (13), probably represents the more favorable mechanism for alleviating steric strain (12). Consistent with this conclusion, in the structure of the RB69 polymerase involving a template containing the cis-(5'R,6'S) Tg lesion and an incoming dATP, the Tg CH\textsubscript{3} group is in the axial conformation, despite hindering stacking of the adjacent 5'-template guanine (71). Bolton and co-workers (11) reported that a disordered structure resulted when 5'R-Tg was placed into a duplex containing the 5'-GTgC-3' sequence. This suggests that formation of an intra-strand hydrogen between the Tg C6 OH and the N7 position of a 3' purine (13) is important in stabilizing the cis-(5'R,6'S) Tg lesion in duplex DNA. The observation that the \( K_d \) values of XPC/HR23B from either the 5'R-Tg- (~18 nM) or AAF-DNA substrates (~27 nM) are lower than those from XPA is consistent with the proposed role of XPC/HR23B as the damage recognition factor in human genomic NER (64,69). In comparison, XPC/HR23B binds to cisplatin 1,3-intrastrand adducts or 6-nt lengths of mispaired DNA with a \( K_d \) of 1–3 nM (64). The recognition of the AAF-damaged DNA by XPC/HR23B has been attributed to the adduct's inability to base pair efficiently with cytosine (72), and an emerging consensus posits that
disruption of normal base pairing and the resulting destabilization of the helix, rather than the recognition of helical distortion by bulky lesions, governs the affinity of XPC/HR23B binding (73). It should be noted, however, that while 5R-Tg can be recognized by XPC/HR23B and XPA, the efficiency of incision of this site-specific lesion by the human NER system remains to be determined.

From studies of the yeast XPC orthologue Rad4 bound to DNA containing a cyclobutane pyrimidine dimer, Min and Pavletich (74) concluded that damage recognized by Rad4 destabilizes the helix and facilitates the flipping out of two base pairs by the protein. The present data are consistent with this model. The presence of 5R-Tg in the duplex perturbs the 5'-neighbor base pair G<sup>5</sup>*C<sup>20</sup>, in addition to the damaged base pair X<sup>6</sup>*A<sup>19</sup>. As indicated in Figure 5, the imino resonance of base pair G<sup>5</sup>*C<sup>20</sup> broadens due to solvent exchange. In fact, it disappears from the <sup>1</sup>H NMR spectrum ~35°C lower in the 5R-Tg modified DNA as compared to the unmodified duplex (12).

This destabilization of the two base pairs by the cis-(5R,6S) epimer may lower the activation barrier with respect to flipping both base pairs out of the helix, enabling XPC/HR23B to recognize the 5R Tg lesion prior to the recruitment of XPA. Based upon MD simulations of the Dickerson-Drew 5R-Tg-modified dodecamer d(CGCGAATg<sub>C15</sub>C15)G<sub>C14</sub>, both Miller et al. (75) and Miaskiewicz et al. (76) also concluded that Tg weakened Watson–Crick hydrogen bonds of the 5'-neighbor base pair. The increase in transverse relaxation for Tg CH<sub>3</sub> as compared to unmodified thymine CH<sub>3</sub> protons is attributed to the puckering of the Tg six-member ring. However, increased backbone and sugar motions in tandem with the puckering of the Tg ring cannot be excluded.

The 5R-Tg lesion is also a substrate for base excision repair in <i>E. coli</i> and in mammalian cells (77). In <i>E. coli</i>, repair of Tg is initiated by endonuclease III (Nth) (78) and endonuclease VIII (Nei) (79). Yeast (80), mammalian (81,82) and human orthologs (83–85) of Nth have been characterized. Likewise, human orthologs of Nei have been characterized (86,87). The human hNTH1 exhibits a 13:1 preference for excising the 5R versus the 5S epimers, whereas hNEIL1 (86,88) shows a 1.5:1 preference for excising the 5R versus the 5S epimers (89). Similar observations have been made for prokaryotic, yeast and murine glycosylases (90). Ocampa-Hafalla et al. (17) observed that the repair of Tg by DNA N-glycosylases/AP lyases is modulated by the cis–trans epimerization of these two sets of diastereomers and that repair of 5R Tg by hNEIL1 depends upon the opposing base, with Tg•G pairs being excised much more rapidly than Tg•A pairs. Significantly, when 5-methyl cytosine is oxidatively damaged, forming 5-methylcytosine glycol, hydrolytic deamination yields Tg mismatched with dG in (3). Computational studies suggest that substrates for hNEIL1 possess in common a pyrimidine-like ring and hydrogen bond donor-acceptor properties, allowing them to be accommodated within the enzyme’s binding pocket (91). Based upon structural data for the 5R-Tg lesion in a T•G mismatch, we suggested that the wobble orientation of the cis-(5R,6S) Tg base in the Tg•G pair shifts it toward the major groove, reflected in an increased solvent accessible surface and reduced barrier for intrinsic breathing of the lesion (14). While the Tg•A structure also shows an increased solvent accessible surface area, it differs from the Tg•G pair in that an efficient hydrogen bond between Tg OH6→G<sup>3</sup>N7 is possible; this may be sufficient to reduce the barrier toward base-flipping of the Tg lesion into the active site pocket of the glycosylase.

In summary, in the 5'-GTG-3' sequence, 5R-Tg, paired with dA, is a good binding substrate for the NER proteins from <i>E. coli</i> and humans, corroborating earlier reports (18–20). It is a better substrate of the human NER proteins XPA and XPC/HR23B in comparison to the C8 dG adduct of AAF. The structure of the cis-(5R,6S) Tg lesion in this same sequence is consistent with the repair data in that the lesion causes destabilization of the 5'-neighboring GC base pair, in addition to an increase in the solvent accessible area of Tg. It is conceivable that these thermodynamic perturbations of two base pairs contribute toward efficient recognition and NER of the 5R-Tg lesion.

ACCESSION NUMBERS
PDB ID 2KH5, PDB ID 2KH6.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
Dr Markus Voehler and Dr Don Stec assisted with NMR spectroscopy. Dr Nicholas Uylanov and Dr Jarrod A. Smith assisted with structural refinement. Professor D. A. Johnson, ETSU, provided the human serum albumin. Drs Thomas M. and Constance M. Harris provided constructive comments.

FUNDING
National Institutes of Health Grants R01 CA-055678 and P30 ES-000267 (to M.P.S.) and R01 ES-013324 (to A.K.B). Funding for open access charge: National Institutes of Health Grant R01 CA-55678.

Conflict of interest statement. None declared.

REFERENCES
1. Teoule,R., Bonicel,A., Bert,C., Cadet,J. and Polverelli,M. (1974) Identification of radioproducts resulting from the breakage of thymine moiety by gamma irradiation of E. coli DNA in an aerated aqueous solution. Radiat. Res., 57, 46–58.
2. Frenkel,K., Goldstein,M.S. and Teebor,G.W. (1981) Identification of the cis-thymine glycol moiety in chemically oxidized and gamma-irradiated deoxyribonucleic acid by high-pressure liquid chromatography analysis. Biochemistry, 20, 7566–7571.
3. Zuo,S., Boorstein,R.J. and Teebor,G.W. (1995) Oxidative damage to 5-methylcytosine in DNA. Nucleic Acids Res., 23, 3239–3243.
4. Pfeifer, G.P. (2000) p53 mutational spectra and the role of methylated CpG sequences. *Mutat. Res.*, 450, 155–166.

5. Vaishnav, Y., Holwitt, E., Swenberg, C., Lee, H.C. and Kan, L.S. (1991) Synthesis and characterization of stereoisomers of 5,6-dihydro-5,6-dihydroxy-thymidine. *J. Biol. Mol. Struct. Dyn.*, 8, 935–951.

6. Lustig, M.J., Cadet, J., Boorstein, R.J. and Teebor, G.W. (1992) Synthesis of the stereoisomers of thymine glycol, determination of concentrations and rates of interconversion of their cis-trans epimers at equilibrium and demonstration of differential alkali lability within DNA. *Nucleic Acids Res.*, 20, 4839–4845.

7. Wang, Y. (2002) HPLC isolation and mass spectrometric characterization of two isomers of thymine glycols in oligodeoxynucleotides. *Chem. Res. Toxicol.*, 15, 671–676.

8. Cathecart, R., Schwiers, E., Saul, R.L. and Ames, B.N. (1984) Thymine glycol and thymidine glycol in human and rat urine: a possible assay for oxidative DNA damage. *Proc. Natl Acad. Sci. USA*, 81, 5633–5637.

9. Adelman, R., Saul, R.L. and Ames, B.N. (1988) Oxidative damage to DNA: relation to species metabolic rate and life span. *Proc. Natl Acad. Sci. USA*, 85, 2706–2708.

10. Kung, H.C. and Bolton, P.H. (1997) Structure of a duplex DNA containing a thymine glycol residue in solution. *J. Biol. Chem.*, 272, 9227–9236.

11. Kao, J.Y., Goljer, I., Phan, T.A. and Bolton, P.H. (1993) Characterization of the effects of a thymine glycol residue on the structure, dynamics, and stability of duplex DNA by NMR. *J. Biol. Chem.*, 268, 17787–17793.

12. Brown, K.L., Adams, T., Jasti, V.P., Basu, A.K. and Stone, M.P. (2008) Interconversion of the cis-5R,6S- and trans-5R,6R-thymine glycol lesions in duplex DNA. *J. Am. Chem. Soc.*, 129, 11701–11710.

13. Clark, J.M., Pattabiraman, N., Jarvis, W. and Beardsley, G.P. (1987) Modeling and molecular mechanical studies of the cis-thymine glycol radiation damage lesion in DNA. *Biochemistry*, 26, 5404–5409.

14. Brown, K.L., Basu, A.K. and Stone, M.P. The cis-(5R,6S)-thymine glycol lesion occupies the wobble position when mismatched with dG in DNA. *Biochemistry*, 48, 9722–9733.

15. Ide, H., Kow, Y.W. and Wallace, S.S. (1985) Thymine glycols and urea residues in M13 DNA constitute replicative blocks in vitro. *Nucleic Acids Res.*, 13, 8035–8052.

16. Clark, J.M. and Beardsley, G.P. (1986) Thymine glycol lesions terminate chain elongation by DNA polymerase I in vitro. *Nucleic Acids Res.*, 14, 737–749.

17. Vold, R.L., Waugh, J.S., Klein, M.P. and Phelps, D.E. (1968) Measurement of spin relaxation in complex systems. *J. Chem. Phys.*, 58, 3831–3832.

18. Goddard, T.D. and Kneller, D.G. (2006) SPARKEY v. 3.113. University of California, San Francisco.

19. Keepers, J.W. and James, T.L. (1984) A theoretical study of distance determinations from NMR – two-dimensional nuclear Overhauser effect spectra. *J. Magn. Reson.*, 87, 404–426.

20. James, T.L. (1991) Relaxation matrix analysis of two-dimensional nuclear Overhauser effect spectra. *Curr. Opin. Struct. Biol.*, 1, 1042–1053.

21. Borgias, B.A. and James, T.L. (1989) Two-dimensional nuclear Overhauser effect: complete relaxation matrix analysis. *Methods Enzymol.*, 176, 169–183.

22. Borgias, B.A. and James, T.L. (1990) MARDIGRAS – a procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *J. Magn. Reson.*, 87, 475–487.

23. Liu, H., Spielmann, H.P., Ulyanov, N.B., Wemmer, D.E. and James, T.L. (1995) Intermolecular distance bounds from 2D NOE intensities: effect of experimental noise and peak integration errors. *J. Biomol. NMR*, 6, 390–402.

24. Delaglio, F., Wu, Z. and Bax, A. (2001) Measurement of homonuclear proton couplings from regular 2D COSY spectra. *J. Magn. Reson.*, 149, 276–281.

25. Vold, R.L., Waugh, J.S., Klein, M.P. and Phelps, D.E. (1968) Thymine glycol and nuclear magnetic resonance. *Eur. J. Biochem.*, 178, 1–38.

26. van Wijk, J.W., Huckriede, B.D., Ippe, J.H. and Altona, C. (1992) Furanose sugar correlations in DNA from NMR coupling constants. *Methods Enzymol.*, 211, 286–306.

27. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J.A., Vreven, T., Kudin, K.N., Burant, J.C. et al. (2004) GAUSSIAN 03, Gaussian, Inc., Wallingford, CT.

28. Case, D.A., Cheatham, T.E. III, Darden, T., Gohlke, H., Luo, R., Merz, K.M. Jr, Onufriev, A., Simmerling, C., Wang, B. and Woods, R.J. (2005) The AMBER biomolecular simulation programs. *J. Comput. Chem.*, 26, 1688–1688.

29. Cieslik, T. and Case, D.A. (1998) In Leontes, N.B. and SantaLucia, J. Jr (eds), *Molecular Modeling of Nucleic Acids*, Vol. 682. American Chemical Society, Washington, D.C.

30. Wang, J.M. and Case, D.A. (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers*, 56, 275–291.

31. Bashford, D. and Case, D.A. (2000) Generalized Born models of macromolecular solution effects. *Annu. Rev. Phys. Chem.*, 51, 129–152.

32. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Phys. Chem.*, 88, 4364–4367.

33. Langer, R.N. and Lifson, A. (1999) Langevin dynamics of peptides: the frictional dependence of alpha carbon, amide group and side chain backbone dihedral angles. *J. Comp. Chem.*, 20, 1049–1074.

34. Cieslik, T. and Case, D.A. (1998) In Leontes, N.B. and SantaLucia, J. Jr (eds), *Molecular Modeling of Nucleic Acids*, Vol. 682. American Chemical Society, Washington, D.C.

35. Case, D.A., Cheatham, T.E. III, Darden, T., Gohlke, H., Luo, R., Merz, K.M. Jr, Onufriev, A., Simmerling, C., Wang, B. and Woods, R.J. (2005) The AMBER biomolecular simulation programs. *J. Comput. Chem.*, 26, 1688–1688.

36. Cieslik, T. and Case, D.A. (1998) In Leontes, N.B. and SantaLucia, J. Jr (eds), *Molecular Modeling of Nucleic Acids*, Vol. 682. American Chemical Society, Washington, D.C.

37. Cieslik, T. and Case, D.A. (1998) In Leontes, N.B. and SantaLucia, J. Jr (eds), *Molecular Modeling of Nucleic Acids*, Vol. 682. American Chemical Society, Washington, D.C.

38. Wang, J.M. and Case, D.A. (2000) Generalized Born models of macromolecular solution effects. *Annu. Rev. Phys. Chem.*, 51, 129–152.

39. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Phys. Chem.*, 88, 4364–4367.

40. Langer, R.N. and Lifson, A. (1999) Langevin dynamics of peptides: the frictional dependence of alpha carbon, amide group and side chain backbone dihedral angles. *J. Comp. Chem.*, 20, 1049–1074.

41. Bashford, D. and Case, D.A. (2000) Generalized Born models of macromolecular solution effects. *Annu. Rev. Phys. Chem.*, 51, 129–152.

42. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Phys. Chem.*, 88, 4364–4367.
66. Buterin, T., Meyer, C., Giese, B. and Naegeli, H. (2005) DNA quality
62. Gunz, D., Hess, M.T. and Naegeli, H. (1996) Recognition of DNA
64. Hey, T., Lipps, G., Sugasawa, K., Iwai, S., Hanaoka, F. and Krauss, G.
53. Meneni, S.R., D’Mello, R., Norigian, G., Baker, G., Gao, L.,
50. Sanner, M.F., Olson, A.J. and Spehner, J.C. (1996) Reduced surface:
60. Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M.,
57. Cleaver, J.E. (2005) Cancer in xeroderma pigmentosum and related
52. Zou, Y., Shell, S.M., Utzat, C.D., Luo, C., Yang, Z., Geacintov, N.E.,
48. Ravishanker, G., Swaminathan, S., Beveridge, D.L., Lavery, R. and
35. Chiarelli, M.P. and Cho, B.P. (2006) Sequence effects of
41. Yang, Z.G., Liu, Y., Mao, L.Y., Zhang, J.T. and Zou, Y. (2002)
47. Lavery, R. and Sklenar, H. (1988) The definition of generalized
48. Ravishanker, G., Swaminathan, S., Beveridge, D.L., Lavery, R. and
59. Volker, M., Mone, M.J., Karmakar, P., van Hoffen, A., Schul, W.,
59. Volker, M., Mone, M.J., Karmakar, P., van Hoffen, A., Schul, W.,
46. Hey, T., Lipps, G., Sugasawa, K., Iwai, S., Hanaoka, F. and Krauss, G.
65. Dip, R., Camenisch, U. and Naegeli, H. (2004) Mechanisms of DNA
45. Aspinwall, R., Rothwell, D.G., Roldan-Arjona, T., Anselmino, C.,
34. Buterin, T., Meyer, C., Giese, B. and Naegeli, H. (2005) DNA quality
43. Buterin, T., Hess, M.T., Luneva, N., Geacintov, N.E., Amin, S.,
42. Giese, B. and Naegeli, H. (2005) DNA quality control by conformational readout on the undamaged strand of the double helix. Chem. Biol., 12, 913–922.
41. Yang, Z.G., Liu, Y., Mao, L.Y., Zhang, J.T. and Zou, Y. (2002) Dimerization of human XPA and formation of XPA2-RPA protein complex. Biochemistry, 41, 13012–13020.
38. Yang, Z., Roginskaya, M., Colis, L.C., Basu, A.K., Shell, S.M.,
37. Buterin, T., Meyer, C., Giese, B. and Naegeli, H. (2005) DNA quality
34. Buterin, T., Hess, M.T., Luneva, N., Geacintov, N.E., Amin, S.,
endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue. *J. Biol. Chem.*, 273, 21585–21593.

86. Hazra, T.K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y.W., Jaruga, P., Dizdaroglu, M. and Mitra, S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl Acad. Sci. USA*, 99, 3523–3528.

87. Hazra, T.K., Kow, Y.W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkapati, S.K., Mitra, S. and Izumi, T. (2002) Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *J. Biol. Chem.*, 277, 30417–30420.

88. Bandaru, V., Sunkara, S., Wallace, S.S. and Bond, J.P. (2002) A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair (Amst)*, 1, 517–529.

89. Katafuchi, A., Nakano, T., Masaoka, A., Terato, H., Iwai, S., Hanaoka, F. and Ide, H. (2004) Differential specificity of human and *Escherichia coli* endonuclease III and VIII homologues for oxidative base lesions. *J. Biol. Chem.*, 279, 14464–14471.

90. Miller, H., Fernandes, A.S., Zaika, E., McTigue, M.M., Torres, M.C., Wente, M., Iden, C.R. and Grollman, A.P. (2004) Stereoselective excision of thymine glycol from oxidatively damaged DNA. *Nucleic Acids Res.*, 32, 338–345.

91. Jia, L., Shafirovich, V., Geacintov, N.E. and Broyde, S. (2007) Lesion specificity in the base excision repair enzyme hNeil1: modeling and dynamics studies. *Biochemistry*, 46, 5305–5314.