Structure, Flexibility, and Repair of Two Different Orientations of the Same Alkyl Interstrand DNA Cross-Link†

David M. Noll,‡ Mateus Webba da Silva,§ Anne M. Noronha,|| Christopher J. Wilds,|| O. Michael Colvin,§ Michael P. Gamcsik,§ and Paul S. Miller*,⊥

Department of Biophysics and Biophysical Chemistry, School of Medicine, Johns Hopkins University, 725 North Wolfe Street, Baltimore, Maryland 21205, Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710, Department of Chemistry and Biochemistry, Concordia University, Montreal, Quebec H4B 1R6, Canada, and Department of Biochemistry and Molecular Biology, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, Maryland 21205

Received January 4, 2005; Revised Manuscript Received March 8, 2005

ABSTRACT: Interstrand DNA cross-links are the principal cytotoxic lesions produced by chemotherapeutic bifunctional alkylating agents. Using an N^4C−ethyl−N^4C interstrand DNA cross-link to mimic this class of clinically important cancer chemotherapeutic agents, we have characterized the repair, structure, and flexibility of DNA that contains this cross-link in two different orientations. Plasmid DNAs in which the cytosines of single CpG or GpC steps are covalently linked were efficiently processed by repair proficient and homologous recombination deficient strains of *Escherichia coli.* Repair in a nucleotide excision repair (NER) deficient strain was less efficient overall and displayed a 4-fold difference between the two cross-link orientations. Both the structure and flexibility of DNA containing these cross-links were examined using a combination of ^1H NMR, restrained molecular dynamics simulations, and atomic force microscopy (AFM). The NMR structure of a decamer containing a CpG interstrand cross-link shows the cross-link easily accommodated within the duplex with no disruption of hydrogen bonding and only minor perturbations of helical parameters. In contrast, disruptions caused by the GpC cross-link produced considerable conformational flexibility that precluded structure determination by NMR. AFM imaging of cross-link-containing plasmid DNA showed that the increased flexibility observed in the GpC cross-link persists when it is embedded into much larger DNA fragments. These differences may account for the different repair efficiencies seen in NER deficient cells.

DNA cross-links that covalently join both strands of the helix arrest replication, block transcription, and induce sister chromatid exchange (*1*−*4*). These interstrand cross-links are highly cytotoxic and are believed to be responsible for the therapeutic efficacy of a number of clinically important compounds, including the nitrogen mustards (melphalan, chlorambucil, cyclophosphamide, and ifosfamide), alkyl alkane sulfonate (busulfan), chloroethylnitrosoureas (carmustine and lomustine), platinum drugs (cisplatin and carboplatin), and the natural product mitomycin C (*5*). Endogenous cellular processes such as the peroxidation of lipids can also produce interstrand DNA cross-links through the formation of unsaturated aldehydes such as acrolein and crotonaldehyde (*6*). There is considerable evidence that cells have developed mechanisms capable of sensing the presence of interstrand DNA cross-links and that this recognition can lead to the repair of the cross-link or to accelerated cell death (*4*, *7*−*10*). Indeed, the repair and/or tolerance of interstrand DNA cross-links may play a role in the development of tumors resistant to nitrogen mustards (*11*). For example, the resistance of a medulloblastoma cell line to cyclophosphamide and other nitrogen mustards was shown to be associated with the ability of cells to remove interstrand cross-links (*12*, *13*).

The repair of interstrand DNA cross-links is complicated by the absence of a lesion-free strand serving as an undamaged template for DNA replication as well as by the physical constraints placed upon the helix (i.e., preventing strand separation) by the covalent attachment of the two strands. A number of DNA repair processes, including homologous recombination and nucleotide excision repair (NER),[1] are known to play a role in repairing these lesions; however, a clear understanding of the mechanisms that are involved is still lacking. The incision/recombination model originally proposed by Cole is most often cited as being

† This research was supported by a grant from the National Cancer Institute (CA082785). A.M.N. and C.J.W. were each supported in part by postdoctoral fellowships from the Natural Sciences and Engineering Research Council of Canada (NSERC). Duke University NMR Center was established with grants from the NIH, the NSF, and the North Carolina Biotechnology Center.

* To whom correspondence should be addressed. E-mail: pmiller@jhsph.edu. Phone: (410) 955-3489. Fax: (410) 955-2926.

‡ School of Medicine, Johns Hopkins University.

§ Duke University Medical Center.

|| Concordia University.

⊥ Bloomberg School of Public Health, Johns Hopkins University.

[1] Abbreviations: NER, nucleotide excision repair; AFM, atomic force microscopy; NMR, nuclear magnetic resonance; HCR, host cell reactivation; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; COSY, two-dimensional correlated spectroscopy; DQF-COSY, double-quantum-filtered COSY; WLC, wormlike chain; FAPY, formamidopyrimidine; rMD, restrained molecular dynamics.

10.1021/bi050014n CCC: $30.25
Published on Web 04/12/2005

responsible for the repair of interstrand cross-links (*14*, *15*). This model was proposed to account for the requirement of the *uvrA*, *uvrB*, *uvrC*, *uvrD*, *recA*, and *polA* gene products for the in vivo survival of *Escherichia coli* in the presence of the photoreactive cross-linker, psoralen. The biochemical role of many of the components implicated by Cole has been demonstrated in vitro using purified proteins and model DNA substrates. The first step in the bacterial repair of interstrand DNA cross-links is incision by the UvrABC exinuclease both 5′ and 3′ of the cross-link (*16*, *17*). The exact location of the incisions depends on the type of interstrand cross-link present (*16*, *18*); however, incision at psoralen interstrand cross-links, the most thoroughly characterized, occurs at the ninth phosphodiester bond 5′ and the third phosphodiester 3′ of the furan-side thymine of the cross-link (*16*, *17*). Following the dual incisions, the 5′-exonuclease activity of DNA polymerase I can create a single-stranded gap at the incised cross-link (*19*). This single-stranded region can serve as an initiation site for RecA-dependent strand invasion of a homologous undamaged DNA strand (*19*, *20*). This repair intermediate serves as a substrate for the UvrABC exinuclease incision of the remaining pyrone-side thymine adduct in a manner analogous to normal NER repair (*19*, *20*). NER-dependent DNA synthesis followed by ligation results in the error-free repair of interstrand DNA cross-links.

In addition to this error-free pathway, Sladek and coworkers demonstrated that repair can occur in the absence of an undamaged template (*21*). This repair process proceeded via an error-prone pathway that produces point mutations when the bacteria were constitutively induced for the SOS response and produced large deletions when the SOS response was not induced (*21*). Loechler and coworkers, using plasmids carrying a single defined cross-link, also found evidence of a recombination-independent repair pathway (*22*). In subsequent work, a role for DNA polymerase II (*polB*) was demonstrated and a damage tolerance mechanism utilizing translesion synthesis was proposed (*23*).

Repair of interstrand DNA cross-links in eukaryotic cells appears to be more complicated and is less well characterized than prokaryotic repair. Repair in *Saccharomyces cerevisiae* involves one or more of the following pathways: nucleotide excision repair (*RAD3*), homologous recombination (*RAD52*), and postreplication/translesion repair (*RAD6* and *RAD18*) (*9*). Studies in cell extracts and living cells clearly show that mammalian cells are capable of repairing interstrand cross-links. Mammalian NER genes *ERCC1* and *XPF* and homologous recombination genes *XRCC2* and *XRCC3* appear to play important roles in this process (*24*). Mechanistic studies using purified components of the mammalian NER pathway and a psoralen interstrand cross-link resulted in dual incisions, both of which occur 5′ of the cross-linked base (*25*). This is in contrast to the incisions that flank the cross-link produced by the bacterial NER components. The resultant gap leads to "futile" DNA synthesis which may in turn serve as a signal for other cellular responses to interstrand DNA cross-links (*26*). Recent studies have also shown that an error-prone pathway involving all of the known NER genes is involved in psoralen cross-link repair (*27*).

Compared to repair of bulky lesions and intrastrand cross-links, relatively little is known about the mechanism(s) by which interstrand cross-links are recognized and repaired.

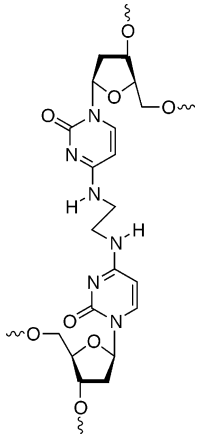

FIGURE 1: Structure of the N⁴C−ethyl−N⁴C cross-link.

One limitation to such studies has been the inability to prepare large quantities of DNA duplexes that contain chemically defined interstrand cross-links (*28*). We have developed automated techniques for chemically synthesizing short DNA duplexes that contain a single N⁴C−alkyl−N⁴C interstrand cross-link (Figure 1) (*29*, *30*). Through variation of the orientation (i.e., sequential positioning) and/or the chain length of the alkyl chain, DNA duplexes have been prepared that mimic therapeutically relevant cross-links and that have unique physical and structural properties. In this report, we examine the correlation between repair and the structural and conformational perturbations that result when an N⁴C−ethyl−N⁴C interstrand cross-link is placed in a CpG and a GpC step.

MATERIALS AND METHODS

Preparation of Cross-Linked DNA. Cross-link-containing oligomers were synthesized and purified using methods previously described to prepare short DNA duplexes that contain the N⁴C−ethyl−N⁴C interstrand cross-link (*31*). Preparation of the cross-link-containing plasmid DNA is described below. A plasmid, pICC2, was prepared by introducing two BbsI recognition sequences into pUC19 such that digestion produces a pair of 5′-GGGG overhangs. Enzymatically phosphorylated cross-linked duplexes containing 5′-CCCC overhangs were incubated with linearized pICC2 in the presence of T4 DNA ligase. Closed circular plasmid DNA was purified following electrophoresis in the presence of saturating amounts of ethidium bromide. Closed circular plasmid DNA prepared in this manner was quantified by comparison with known quantities of plasmid DNA (Figure S1 of the Supporting Information). Concentrations for closed circular plasmids were typically 3−6 ng/μL, and the overall recovery of plasmid DNA was ~50 ng. A canonical DNA duplex comprised of d-CCCCAAAACTTTT and d-TTTTGAAAACCCC strands contains a central G·C base pair and was used as a non-cross-linked control. The quality and purity of these plasmid preparations were assayed using a variety of techniques. (a) Atomic force microscopy demonstrated that more than 95% of the plasmid DNAs were circular and less than 2% existed as concatamers (Figure S2 of the Supporting Information). (b) Native and denaturing gel electrophoresis of radiolabeled plasmid DNA revealed more than 99.5% of the plasmid prepared in this manner contained the cross-link (Figure S3 of the Supporting

Information). (c) A PCR-based contamination assay was used to estimate a conservative limit of less than 0.5% possible contamination with a non-cross-link-containing plasmid.

Cross-Link Repair Assay. Repair of N^4C-ethyl$-N^4C$ cross-link-containing plasmid DNA was assessed using a host cell reactivation assay (*22*, *23*). Competent cells of *E. coli* strains FC40, FC348, and PFB50 were transformed via electroporation with 0.5 ng of plasmid DNA containing either the G·C non-cross-linked control or the CpG or GpC cross-linked plasmid. Immediately following electroporation, 20 and 200 μL of the transformation mixture were plated on LB agar plates containing ampicillin and incubated overnight at 37 °C. The resulting ampicillin resistant colonies were counted. The largest source of systematic error in this assay results from the quantification of input DNA. To minimize this, three independent preparations of non-cross-linked and cross-linked plasmids were prepared. The data reported here are the average of the three individual plasmid preparations. Relative repair efficiencies were determined by dividing the average number of colonies derived from the cross-link-containing plasmids by the number of colonies arising from transformation with the G·C non-cross-linked control plasmid. Electroporation conditions, percent survival, transformation efficiencies, and background contamination levels for each of the *E. coli* strains used are available in Table S1 of the Supporting Information. Sequencing of the repaired plasmids was performed following Qiagen purification of a 5 mL overnight culture inoculated from HCR plates. The DNA duplex used as the G·C non-cross-linked control can insert such that either a G or a C is found in the sequenced strand, and the expected 1:1 ratio of G to C was observed in plasmid DNA arising from the G·C non-cross-link transformation.

NMR Spectroscopy. Standard NMR experiments were carried out on Varian Inova 600 and 800 MHz spectrometers. We recorded NOESY (200 and 60 ms) spectra in ^1H_2O at 0 °C and NOESY (*32*) (40, 60, 100, 160, 250, and 600 ms), DQF-COSY (*33*), and TOCSY (*34*) (spin-lock time of 30 ms) spectra in ^2H_2O at 15 °C. In ^1H_2O data were acquired with a jump-and-return pulse sequence (*35*) and in ^2H_2O with Watergate suppression of the residual water signal (*36*). All data sets were acquired in a phase sensitive mode (TPPI). In ^1H_2O, NOESY data sets were collected with 8K complex points over a spectral width of 16 kHz with 300 t_1 increments. In ^2H_2O, NOESY data sets were collected with 3412 complex points and 300 t_1 free induction decays covering a spectral width of 8 kHz. TOCSY spectra were recorded with 3798 t_2 complex points and 256 t_1 increments, and the DQF-COSY spectrum was acquired with 3468 t_2 complex points over 512 t_1 increments. Proton chemical shifts were referenced to internal sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). All data sets were processed using VNMR (Varian Instruments) and FELIX 2000 (Accelrys Inc.). Structures were visualized and figures prepared with Insight II (Accelrys).

Distance and Torsion Restraints. NOE cross-peaks involving exchangeable protons in NOESY spectra (mixing times of 50 and 200 ms) in H_2O buffer were classified as strong (strong intensity at 50 ms), medium (barely observable at 50 ms), and weak (not observable at 50 ms and observable at 200 ms), and the observable proton pairs were restrained to distances of 3.0 ± 0.6, 4.0 ± 1.0, and 5.0 ± 1.2 Å, respectively. NOE buildups for nonexchangeable protons were derived from NOESY spectra in ^2H_2O buffer recorded as a function of mixing time (40, 60, 100, 160, 250, and 600 ms). Distances were estimated from the initial buildup rates within FELIX 2000 (Accelrys). The cytosine H5−H6 interproton distance of 2.46 Å was used as a reference. The upper and lower bounds were allowed to vary ±30%. Overlapping cross-peaks were given generous bounds (up to ±80%). Atoms participating in experimentally identified canonical base pairing (based on NOE patterns) were restrained with distances corresponding to ideal hydrogen bond geometry (*37*) between donor and acceptor atoms.

The appearance of strong H1′−H2″ and very weak to no H2″−H3′ cross-peaks in the DQF-COSY spectrum indicates that the most populated conformations are S-type conformations, with the exception of the tethered oligo C5* and terminal residues. Thus, except for these, δ and the endocyclic $\nu(0)-\nu(4)$ torsion angles were moderately constrained, leaving the sugar free to take any conformation without an energy penalty between C4′-*endo* and O1′-*endo*, including C2′-*endo*. A very strong NOE between the H6 and H3′ protons of C5* (which persists at lower mixing times), together with the observed coupling patterns between the H1′/H3′ protons and H2′ and H2″ protons for the sugar pucker, independently establishes an N-type conformation for this residue, and therefore, loose C3′-*endo* centered restraints were imposed on δ and endocyclic $\nu(0)-\nu(4)$ torsion angles. The phosphorus spectrum occupies a very narrow region (within 1 ppm), indicating an undistorted B-DNA helix of type B_I. Consequently, loose phosphodiester backbone torsion angles were employed for α (60 ± 40°) and ζ (−90 ± 60°) (*38*−*40*). CURVES 5.2 (*41*) was used to estimate DNA conformation and helical parameters.

Distance-Restrained Molecular Dynamics Regularization. Calculations were performed with XPLOR (*42*) using the CHARMm force field (*43*) and adapted for restrained MD for nucleic acids. All calculations were executed in vacuo without explicit counterions. The cross-link was placed in the major groove as defined by the available NMR restraints. The initial distance geometry and simulated annealing refinement protocol started from 250 different structures generated from sets of two strands, each 10 nucleotides long, randomized over all dihedral angles. At this stage of the refinement, planarity restraints were included. A number of structures (34 of 200) emerged separated from nonconverged structures by large gaps in all components of the potential energy function (dihedral angles, van der Waals distances, NOE violations, and covalent geometry). This set was subsequently submitted to further refinement. Sets of rMD calculations were performed using random velocities fitting a Maxwell−Boltzmann distribution. The empirical energy function was developed for nucleic acids and treated all hydrogens explicitly. It consisted of energy terms for hydrogen bonding, and nonbonded interactions, bonds, bond angles, torsion angles, and tetrahedral and planar geometry, including van der Waals and electrostatic forces. The effective function included terms describing distance and dihedral restraints, which were in the form of square well potentials (*44*). Most estimated distances from NOE data analysis were incorporated as ambiguous restraints using the "SUM averaging" option of XPLOR, since they could reflect intrastrand and/or interstrand contributions. On the basis of

the 2-fold symmetry, noncrystallographic symmetry restraints were imposed on all atoms. Planarity restraints were used in all stages of computations. The simulated annealing procedure consisted of a total of 53 ps of rMD, including heating for 7 ps from 300 to 1000 K, a 20 ps scale-up of restraints at high temperatures, cooling for 14 ps to 300 K, and 12 ps of equilibration rMD. The temperature was controlled by coupling the molecules to a temperature bath with a coupling constant of 0.025 ps (*45*). The van der Waals term was approximated using the Lennard-Jones potential energy function, and bond lengths involving hydrogens were fixed with the SHAKE algorithm (*46*) during molecular dynamics calculations. Coordinates of the CpG cross-linked duplex have been deposited in the Protein Data Bank (entry 1N74).

Atomic Force Microscopy. Cross-link-containing or non-cross-linked plasmid DNA (~85 ng) was digested with the restriction enzyme PvuII and purified using a QiaQuick PCR purification kit, and the DNA was eluted with 30 μL of water. The non-cross-linked plasmid DNAs used in the AFM studies are identical in sequence to the cross-linked plasmids, albeit they lack the N⁴C−ethyl−N⁴C cross-link. Samples were adjusted to a final DNA concentration of 0.5−1 ng/μL in a buffer containing 4 mM HEPES (pH 7.4) and 10 mM NaCl. Samples containing DNA (25 μL) were incubated for 2 min at room temperature on freshly cleaved 15 mm disks of ruby muscovite mica in a chamber equilibrated with sample buffer to minimize evaporation. The mica disk was gently rinsed with 10 mL of water and blown dry with compressed gas. Images were acquired using a Nanoscope III controller with a Multimode AFM system (Digital Instruments, Santa Barbara, CA**)**. TAP-300 silicon cantilevers were cleaned by exposure to high-intensity UV radiation prior to use. Samples were imaged by ambient tapping mode AFM at scan speeds of 1−2 Hz; 2.5 μm × 2.5 μm images were obtained (512 × 512 pixels). DNA contour lengths were determined by manually tracing the DNA molecule using ImageJ version 1.3. End-to-end distances were determined by connecting the first and last point of each DNA molecule. Measurements were performed on only molecules that were completely visible in the image and whose path could be unambiguously traced. The significance of the end-to-end measurements for the cross-link and non-cross-linked form was assessed by the Mann−Whitney U test using Abel version 1.5. The *Z*-score for the CpG non-cross-linked (322 molecules) and CpG cross-linked (289 molecules) forms was 0.928; the *Z*-score for the GpC non-cross-linked (296 molecules) and GpC cross-linked (344 molecules) forms was −7.510.

RESULTS

Repair of Cross-Linked Plasmid DNA. Bacterial repair of interstrand DNA cross-links was examined using closed circular plasmid DNA containing a chemically homogeneous interstrand cross-link at a defined position (Figure 2). Synthetic DNA duplexes in which the cytosines of either a CpG or a GpC step are covalently joined via a N⁴C−ethyl−N⁴C cross-link were prepared such that they could be ligated into a specially constructed plasmid (Figure 2a). Closed circular plasmid DNA containing a CpG or a GpC interstrand cross-link was purified from ligation mixtures of cross-linked insert and linearized plasmid DNA (Figure 2b). This procedure exploits the fact that in the presence of saturating

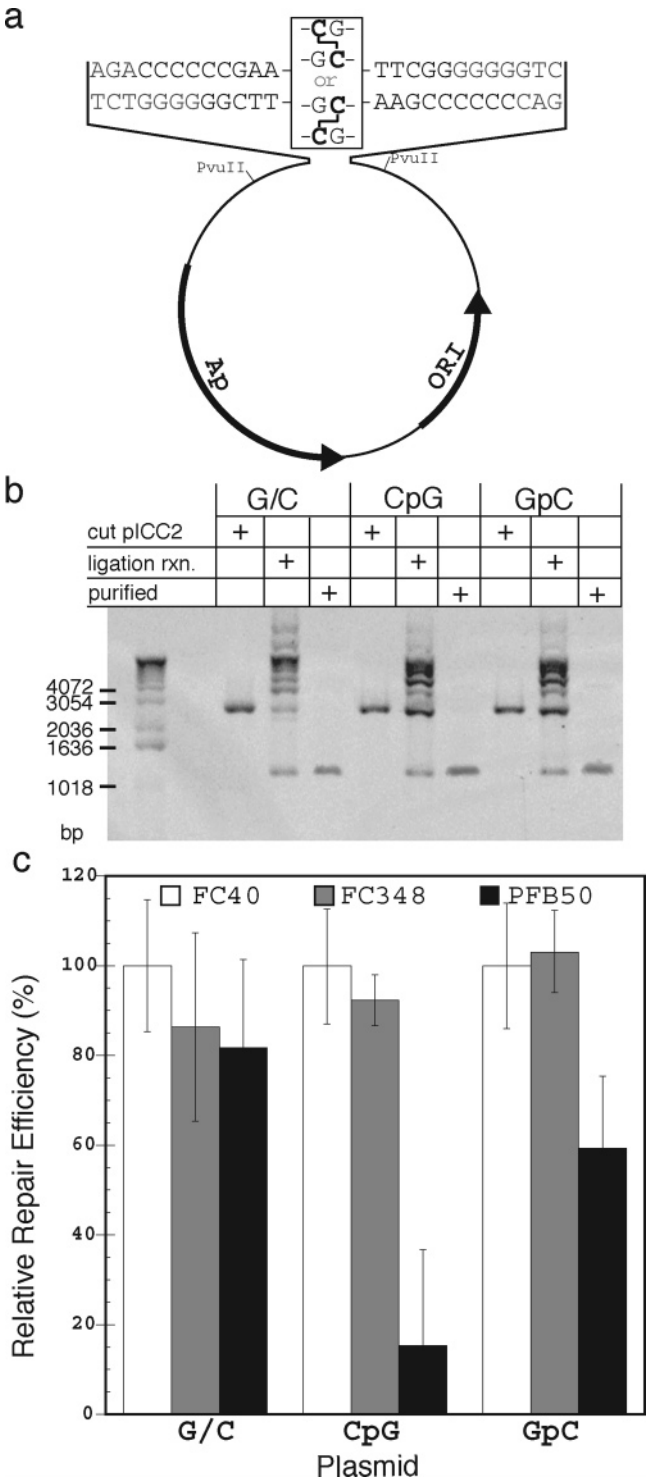

FIGURE 2: Design, construction, and repair of cross-link-containing plasmid DNA. (a) Scheme depicting the cross-link-containing plasmids. Shown are the positions of the selectable marker ampicillin (Ap), the origin of replication (ORI), the sequence of both cross-links, and the plasmid-derived site of insertion. (b) An agarose gel showing the preparation of the non-cross-linked G·C-containing plasmid and the CpG and GpC cross-links. Shown are the linearized pICC2 vector, the crude ligation reaction, and purified closed circular plasmid DNA. (c) Graph showing the relative repair efficiency of the three prepared plasmids in three strains of *E. coli*: FC40, wild type (white bars); FC348, Δ*recA*, (gray bars); and PFB50, *uvrA6* (black bars).

amounts of an intercalating agent (e.g., ethidium bromide or Syber Green I), closed circular DNA becomes highly supercoiled and migrates significantly faster than other

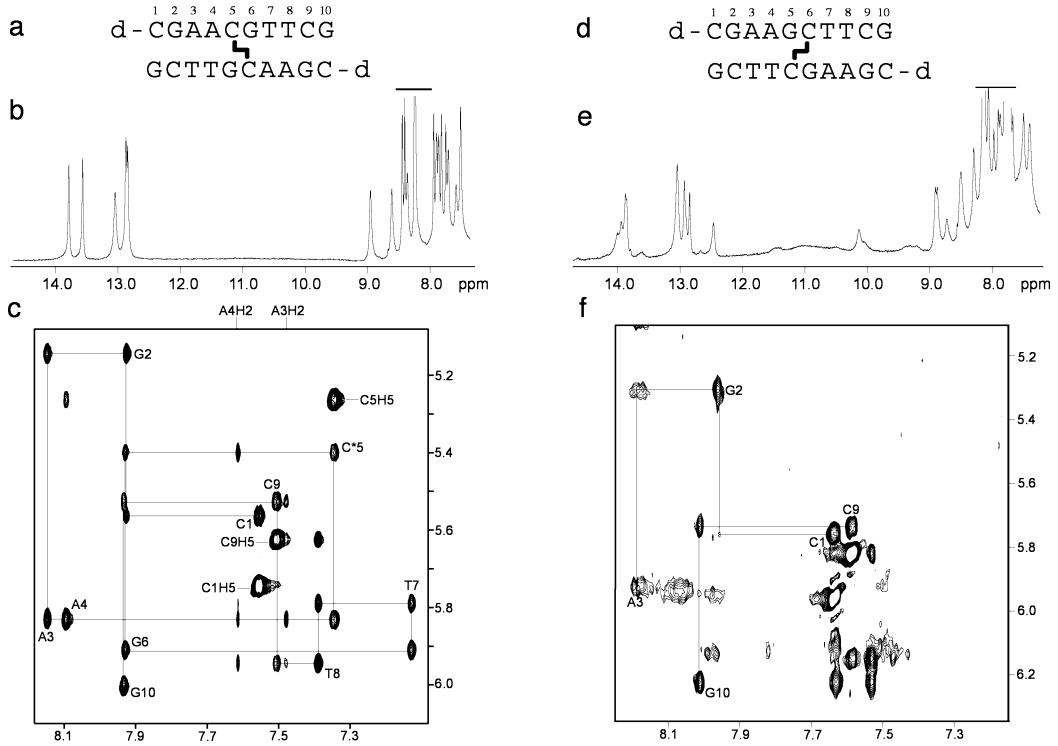

FIGURE 3: NMR structure determination of cross-link-containing oligonucleotides. (a and d) Sequence, cross-link orientation, and numbering of oligonucleotide duplexes used in the NMR studies. (b and e) One-dimensional ^1H NMR spectra of cross-link-containing duplexes in ^1H_2O containing 100 mM NaCl at pH 7.4. (c and f) Selected regions of the NOESY spectra (250 ms) of cross-link-containing duplexes in ^2H_2O at pH 7.4 and 100 mM NaCl at 20 °C showing NOE correlations between H8/H6/H2/H5 and H1′. Where visible, the self-peaks involving H1′ protons have been labeled and their sequential connectivities have been drawn.

components of the ligation reaction mixture. In addition to cross-linked DNA, a canonical DNA duplex containing a central G•C base pair was included as a non-cross-linked control.

The ability of a wild-type *E. coli* strain proficient in known DNA repair processes (FC40), a homologous recombination deficient strain (Δ*recA*) derived from the wild-type strain (FC348), and a NER deficient strain (*uvrA6*) also derived from the wild-type strain (PFB50) to repair cross-link-containing plasmid DNA was determined (Figure 2c). The extent of repair was measured using a host cell reactivation (HCR) assay. This assay provides an aggregate measure of the ability of a particular cell type to replicate the cross-link-containing plasmid. Plasmid DNA that carries both a selectable marker (ampicillin resistance) and an interstrand cross-link is transformed via electroporation into bacterial cells and a relative repair efficiency determined by comparing survival to the non-cross-linked control plasmid. HCR assays have been used to study the bacterial repair of cross-links as well as the cross-link repair in eukaryotic cells (*21*−*23*, *27*, *47*).

The wild-type strain, FC40, efficiently repaired both orientations of the N^4C−ethyl−N^4C cross-link. Similarly, the level of repair in the Δ*recA* strain, FC348, was indistinguishable from wild-type repair levels. The lack of a role for recombination in the repair of these cross-linked plasmids is not surprising given the absence of an undamaged homologous template strand. The presence of either interstrand cross-link did cause significant reductions (approximately 10-fold for CpG and ∼2-fold for GpC) in the repair efficiencies in the NER deficient strain, PFB50. NER is recognized as being one of the pathways responsible for the repair of interstrand cross-linked DNA, and as such, it is not surprising that disruption of the pathway results in an increased sensitivity to such cross-links. Interestingly, the repair of the CpG interstrand cross-link is less efficient than the repair of the GpC interstrand cross-link, suggesting that recognition and/or processing is sensitive to cross-link orientation.

Repaired plasmid DNA was sequenced to examine the fidelity of interstrand cross-link repair. In all cases (approximately 100 plasmids from each cross-link), sequencing revealed that repair of the cross-link occurred with G being placed opposite the N^4C−ethyl−N^4C cytosines. In fact, no mutations, insertions, or deletions were found either at the cross-link site or in the regions surrounding the cross-link.

Determination of the NMR Structure. Previous studies of N^4C−ethyl−N^4C-containing DNA have shown that the CpG cross-link has a significantly higher thermal stability than a duplex containing an GpC interstrand cross-link (*31*), suggesting that the differences observed in the NER deficient bacterial strain may be caused by differential recognition of and/or repair of the cross-link. To examine this possibility, NMR and restrained molecular dynamics simulation were used to study the effect that cross-link orientation has upon the structure and conformation of DNA (Figures 3 and 4). DNA decamers containing the N^4C−ethyl−N^4C cross-link in either the CpG or GpC orientation were prepared (Figure 3a,c). These duplexes, with the exception of the terminal base pairs, are identical in sequence to the cross-linked oligomers used in the repair assay. The one-dimensional imino and aromatic proton spectrum of the CpG cross-linked oligomer shows sharp well-resolved resonances consistent with a single predominant stable structure (Figure 3b), while

the GpC cross-link-containing duplex displays poorly resolved and significantly broadened resonances typically observed when multiple conformations are undergoing exchange during the time scale of the experiment (Figure 3c).

The assignment of resonances in the CpG cross-linked duplex was straightforward. Assignments of the imino and amino exchangeable protons were performed following analysis of 1-1 NOESY spectra with mixing times of 60 and 200 ms and were based on assignments of the nonexchangeable region. NOEs between thymine imino protons and adenine H2 and amino protons across A·T base pairs and between imino and cytosine amino and H5 protons across G·C base pairs are observed (Figure S4 of the Supporting Information). Intrastrand dipolar connectivities such as those for A3 (T8 H3−A3 H61, T8 H3−A3 H62, and T8 H3−A3 H2), A4 (T7 H3−A4 H61, T7 H3−A4 H62, and T7 H3−A4 H2), C1 (G10 H1−C1 H41 and G10 H1−C1 H62), C9 (G2 H1−C9 H41 and G2 H1−C9 H42), and notably cross-linked cytosine C5 (G6 H1−C5* H41) suggest Watson−Crick base pairing throughout the stem. The environment of the alkylated base within the DNA stem is well-characterized through a good number of dipolar connectivities. The well-stacked nature of the consecutive canonical base pairs can be readily traced through several intra- and interstrand cross-peaks: T7 H3−A3 H2, T8 H3−C9 H42, T7 H3−G6 H21, C5* H41−T7 H3, T7 H3−T8 H3, C9 H41−C1 H41, and G2 H1−C1 H41 (Figure S4 of the Supporting Information). The imino and amino chemical shifts at 0 °C have been tabulated and can be found in Table S2 of the Supporting Information.

Assignments of the H8/H6/H5 base and H1′ sugar protons were made through analysis of the 50 and 250 ms NOESY spectra by established methods (*48*, *49*). With a mixing time of 250 ms, we can readily trace the sequential dipolar connectivities between the base and its own 5′-flanking sugar H1′ protons along individual strands as expected for right-handed DNA (Figure 3c). At a low mixing time (50 ms), weak H8/H6−H1′ cross-peaks were observed, indicating that all residues are anti (when compared to C2 H6−C2 H5, C5* H6−C5* H5, and C11 H6−C11 H5 cross-peaks). The C5 H6−A4 H2 cross-peak is significant in establishing the orientation of the tethered base (Figure 3c). The ethyl moiety of the cross-link did not show cross-peaks to sugar pucker protons. This is consistent with its expected location in the major groove.

Cross-strand (i) A H2−N(j+1) H1′ (i and j represent different strands, N is any base) and intrastrand (ii) A H2−N(i+1) H1′ interactions were evident, and we observed contacts among A4 H2−T7 H1′, A4 H2−T8 H1′, and A4 H2−C5* H1′ cross-links, and among A3 H2−T8 H1′, A3 H2−C9 H1′, and A3 H2−A4 H1′ cross-links (Figure 3c). The stereospecific assignments of individual H2′ and H2″ protons were derived from comparison of the intensity of the H1′−H2′ and H1′−H2″ intraresidue cross-peaks in a 50 ms NOESY spectrum. At this mixing time, the intraresidue H1′−H2″ cross-peaks are stronger than the H1′−H2′ cross-peaks in B-DNA. In contrast to all other nucleotides, the chemical shift of G10 H2′ appears at a higher field than that of G10 H2″. DQF-COSY spectral patterns indicated that H1′−H2″ coupling constants were reasonably large (>6 Hz) with only one nonterminal residue (C*5) showing strong H2″−H3′ cross-peaks. Therefore, the sugar geometries are predominantly S-type (*50*), and consequently, the strand structures may be taken to belong to the B-DNA family.

Overall, the N⁴C−ethyl−N⁴C cross-link in the CpG orientation is accommodated by the DNA helix with only minor distortions (Figure 4). The ethyl cross-link protrudes into the major groove and does not interfere with the cross-linked bases participating in canonical Watson−Crick hydrogen bonding interactions. Small distortions caused by the presence of the cross-link include (a) a 1 order of magnitude increase in the base pair propeller twist at the site of the cross-link as compared to the other base pairs in the stem, (b) buckling between base pairs being greatest at the site of the cross-link and decreasing as one moves away from it, (c) *x*-displacement magnitudes intermediate between those of B-DNA (∼0 Å *x*-displacement) and A-DNA (3−4 Å *x*-displacement), (d) the intrastrand interbase twist for the step A4p C5* being dramatically underwound (26°) (most interbase twist steps are close to 36°; the mean magnitude intrastrand interbase twist is 34°), and (e) the cross-linked nucleotide adopting a C4′-*exo* pucker and being perhaps the most notable perturbation caused by the cross-link (the sugar puckers for the remaining stem nucleotides are C2′-*endo*). Inspection of torsion angles ϵ and ζ along the sugar−phosphate backbone reveals the phosphodiester populating predominantly the B_I conformation. The input and structure convergence parameters are listed in Table 1, and geometrical parameters for the solution structure of the CpG cross-linked oligomer in 100 mM NaCl (pH 7.4) are available in Table S3 of the Supporting Information.

In general, these distortions are quite minor and allow the N⁴C−ethyl−N⁴C cross-link in the CpG orientation to be easily accommodated by the DNA helix. In contrast, when the orientation of the cross-link is reversed, the DNA helix undergoes considerable destabilization. The NOESY spectrum of the GpC cross-linked oligomer shows sharp cross-peaks for the residues that comprise the ends of the helix; however, the core portion of the DNA stem undergoes considerable exchange (Figure 3f), and efforts to determine the structure of the GpC-oriented N⁴C−ethyl−N⁴C cross-link were unsuccessful. Thus, the GpC cross-linked bases and the surrounding base pairs appear to be undergoing significant conformational flux.

AFM Measurements of Flexibility. Atomic force microscopy was used to determine if the differences in cross-link structure and stability observed in the short cross-link-containing DNA duplexes persist in the context of larger DNA fragments (Figure 5). Closed circular plasmid DNA containing a CpG or a GpC interstrand cross-link was digested with *Pvu*II to generate two fragments, a 323 bp fragment with the cross-link located 114 bp from one end and a 2.3 kb fragment; both DNA fragments are observed in the AFM images (Figure 5a,b).

DNA was deposited onto mica under conditions that allow the molecule to freely equilibrate in two dimensions before being trapped for imaging. Thus, these DNA molecules represent the average or ensemble properties of the lowest-energy conformation available to the molecule (*51*). The physical properties of DNA deposited under these conditions can be described as long, intrinsically straight polymers using the wormlike chain (WLC) model. This model provides a theoretical underpinning for describing the behavior of DNA

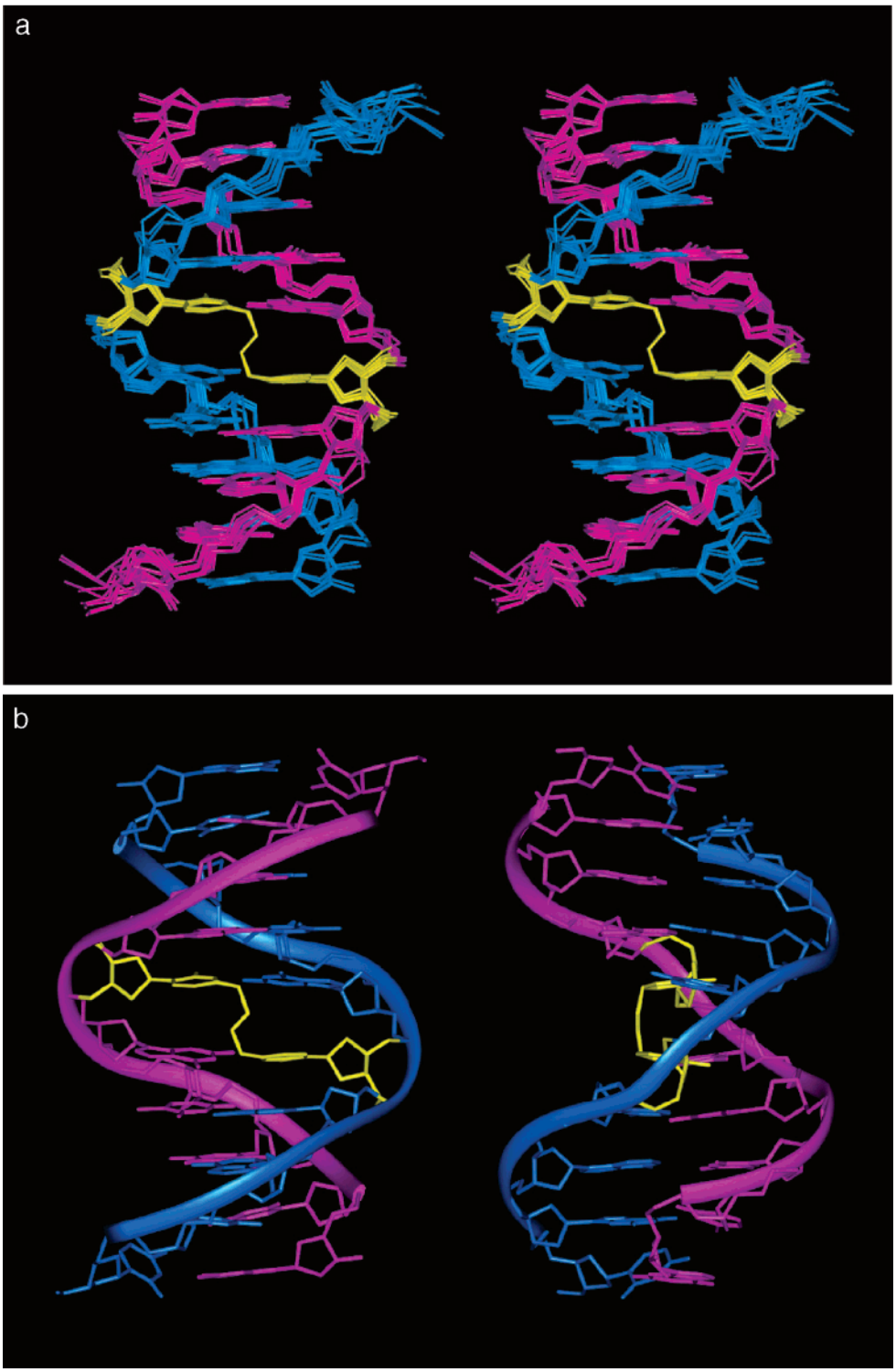

FIGURE 4: Solution structure of the CpG N^4C−ethyl−N^4C interstrand cross-link. (a) Stereoview of a superposition of 10 refined structures of the CpG cross-linked duplex. These structures exhibit average heavy atom root-mean-square deviations (rmsd) from the mean structure of 0.63 Å. (b) Orthogonal views of the duplex structure. The individual strands are colored pink and blue, and the cross-link is colored yellow. Protons and the phosphate oxygen atoms have been omitted for clarity.

molecules observed by AFM and allows biophysical properties such as persistence length and changes therein as well as intrinsic or protein-induced bend angles of DNA molecules to be determined by measuring contour lengths and the end-to-end distances of a large number of DNA molecules (*51*−*53*). The end-to-end distances of a large number of cross-link-containing pieces of DNA and non-cross-linked plasmid DNA whose sequences are identical were determined (Figure 5c,d). The distributions of end-to-end distances for the CpG cross-link and its non-cross-linked counterpart are not significantly different (Z-score = 0.93) and suggest that the CpG interstrand cross-link-containing plasmid DNA is indistinguishable from its non-cross-linked counterpart. In contrast, the end-to-end distances for the GpC interstrand cross-link are smaller than those of its non-cross-linked counterpart (Z-score = −7.51). The introduction of a region

Table 1: Restraints and Refinement Statistics for 10 Selected Structures for the CpG Cross-Linked Oligomer

NMR Distance Restraints[a]	
total no. of restraints	466
no. of nonexchangeable protons	199
no. of exchangeable protons	145
no. of hydrogen bond restraints (empirical)	52
no. of dihedral angle restraints	70
Structural Statistics	
NMR *R*-factor ($R_{1/6}$)	0.105−0.106
total NOE rmsd (Å)	0.028−0.029
no. of NOE violations exceeding 0.2 Å	0
heavy atom pairwise rmsd (Å)	0.8 ± 0.2

[a] Noncrystallographic symmetry restraints on all heavy atoms.

of increased flexibility into an otherwise normal piece of DNA produces a decrease in the end-to-end distance; this behavior has been observed with DNA molecules containing single-stranded gaps (*53*). Thus, it appears that the increased conformational flexibility that was observed in the NMR studies of the short GpC cross-linked decamer persists when this cross-link is embedded in a larger piece of DNA. Such changes in flexibility could be responsible for the differential processing observed in the NER deficient *E. coli* strain.

DISCUSSION

Interstrand DNA cross-links pose a unique set of challenges to the maintenance of genome integrity. These cross-links prevent DNA strand separation and thus block replication and interfere with transcription. It is these properties that are largely believed to contribute to the potent cytotoxicity of the chemotherapeutically important bifunctional alkylating agents. Using the $N^4C-ethyl-N^4C$ alkyl interstrand cross-link as a model, we have studied the repair, structure, and flexibility of cross-linked DNA in two orientations.

The ethyl portion of the $N^4C-ethyl-N^4C$ interstrand cross-link covalently attaches the N^4-exocyclic amino groups of cytosines on opposite strands of the DNA duplex and was predicted to span the major groove. The NMR-derived structure shows that the cross-linked cytosines of the CpG step are positioned such that the ethyl bridge lies along the floor of the major groove and is easily accommodated by the DNA helix (Figure 4**)**. The cross-link produces no significant distortions, and the DNA duplex is essentially unperturbed B-form DNA in which all hydrogen bonding and base stacking interactions are maintained. The structure reveals almost no deviation from idealized helical parameters and does not show any significant alterations in the depth or width of either the major or minor groove. When incorporated into a larger piece of DNA, the CpG cross-link produces a molecule whose physical properties, as measured by AFM, are indistinguishable from those of non-cross-linked DNA.

In contrast to the well-defined helical structure of the CpG cross-linked duplex, the same cross-link in the GpC orientation disrupts the structure of the helix to such an extent that NMR analysis is impossible. When this cross-link is incorporated into plasmid DNA, the distortions evident in the GpC DNA decamer persist and result in a significant decrease in the end-to-end distances of cross-link-containing molecules. The decrease in end-to-end distances is consistent with a significant increase in flexibility at the site of the GpC cross-link in the DNA molecule. This increased flexibility presumably arises from destabilization of the DNA helix due to simple geometric constraints. The ethyl cross-link is significantly shorter than the distance of ~5.6 Å between exocyclic amino groups of the cytosines in a GpC step in B-form DNA. Given the distance constraints, it seems likely that the cross-link leads to the disruption of the base pairing interactions of the cross-linked bases and induces out-of-plane deviations with the loss of intrastrand base stacking. This type of distortion is likely to produce a highly dynamic lesion that produces a localized destabilization extending several base pairs away from the cross-link itself, thus resulting in an overall dynamic lesion site that affects a few base pairs of the DNA stem. Similar geometric constraints are believed to occur with some chemotherapeutic bifunctional alkylating agents. Nitrogen mustards such as cyclophosphamide, melphalan, chlorambucil, and mechlorethamine all produce the same five-atom diethyleneamine tether (7.5 Å) that preferentially cross-links the N^7 atom of guanines in GNC steps (8.9 Å) (*28*). Physical studies of mechlorethamine interstrand cross-linked DNA have confirmed the presence of a cross-link-induced deformation of the helix (*54*).

Previously, we reported the NMR-derived structure of a DNA duplex that contains the $N^4C-ethyl-N^4C$ interstrand cross-link between two directly opposed cytosines (*55*). In this orientation, the cross-link produces an A-form bubble in the central portion of the duplex. This bubble is accompanied by a narrowing of both the major and minor grooves and a concomitant increase in the diameter of the DNA toroid. The two-methylene linker projects into the major groove and results in DNA bending by approximately 27° toward the major groove. The destabilization induced by the cross-link spreads over 3 bp on either side of the cross-link. The deformations caused by the opposed $N^4C-ethyl-N^4C$ interstrand cross-link thus fall between the perturbations of the CpG cross-link which lacks any distortion and those of the GpC cross-link which is so distorted that it fails to adopt any discrete structure but rather appears as an unstructured region that is highly flexible.

The solution structures of two alkyl interstrand cross-links that span the minor groove of DNA have been reported. The structure of a trimethylene interstrand cross-link has been determined in both the CpG and GpC orientations (*56*, *57*). This cross-link was designed to mimic the interstrand cross-link formed by the reaction of DNA with malondialdehyde and consists of a three-methylene linker in which the N^2 exocyclic amino groups of guanine residues are connected (*6*, *58*). The structure of the duplex having the CpG-oriented cross-link was predominately B-form, with pseudoplanar tethered bases. However, the absence of sequential NOEs at the site of the cross-link was consistent with disruption of hydrogen bonds in the immediate vicinity of the cross-link. When the orientation of the guanines was reversed, considerable distortion of the DNA helix occurred. The GpC trimethylene cross-linked duplex resembles neither A- nor B-form DNA, exhibits a dramatic flattening of the major and minor grooves and pseudo-base pair twisting involving the tethered bases, and is associated with a complete abolition of base stacking interactions for the cytosine partners of the cross-linked bases.

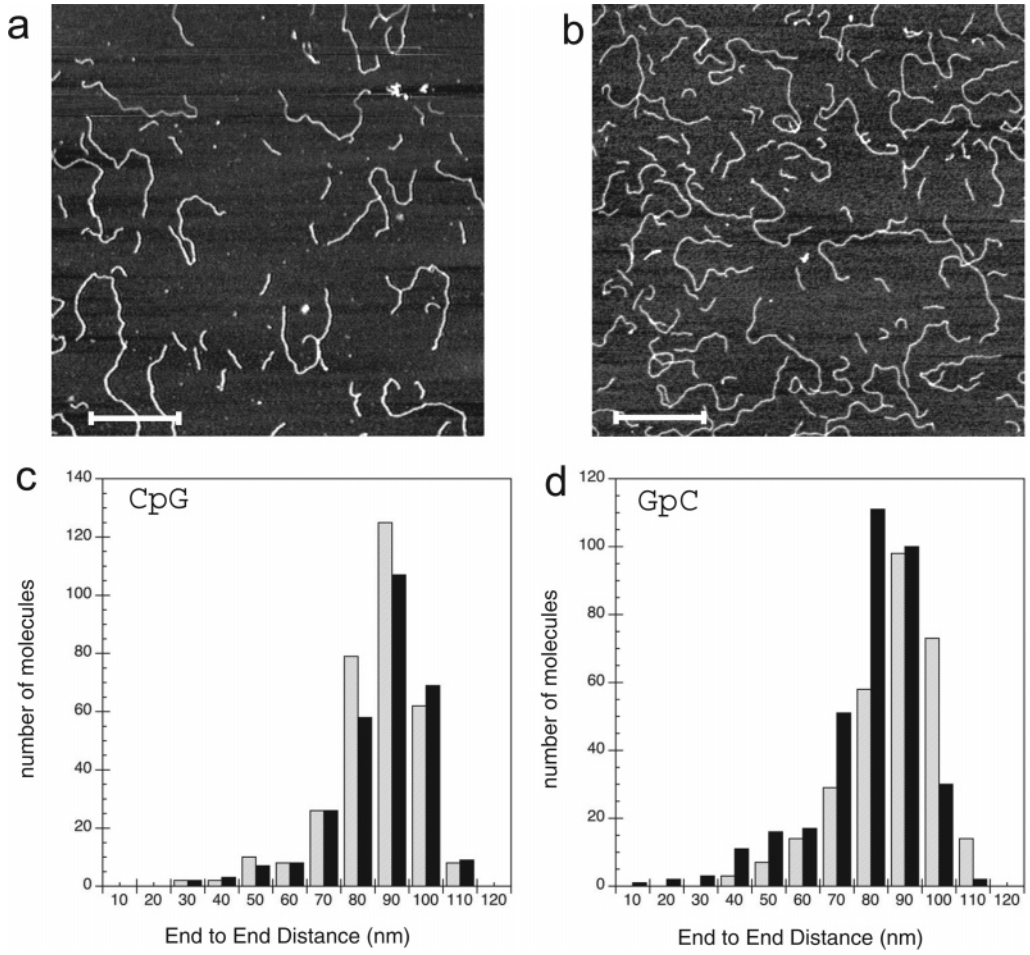

FIGURE 5: Atomic force microscopy of cross-link-containing plasmids. (a and b) AFM images of PvuII-digested CpG and GpC cross-linked plasmid DNA, respectively. The image size is 2.5 μm × 2.5 μm; the white bar is 500 nm. (c and d) Plot of the distribution of end-to-end distances for the cross-linked plasmid (black) and the non-cross-link-containing plasmid (gray) for the CpG and GpC plasmids.

The structure of a nitrous acid interstrand cross-link has also recently been reported (*59*, *60*). Nitrous acid cross-links consist of a single methylene bridge linking the N^2 exocyclic amino groups of the guanines of a CpG step. The structure reveals dramatic perturbations in the DNA structure with the cross-linked guanines forming a nearly planar cross-linked base pair while the cytosine bases opposite the cross-linked guanines are flipped out of the helix and into the minor groove. Additional structures demonstrated that sequence context has a dramatic effect upon the accommodation of the nitrous acid cross-link and the flexibility of the extrahelical cytosines.

Despite advances in the synthesis of alkyl interstrand cross-linked DNA and growing insight into the effect these cross-links have on the structural properties of DNA, our understanding of the pathways and processes responsible for the repair of these lesions is still quite rudimentary. In particular, very little is understood with regard to the interplay of DNA repair processes, including NER, homologous recombination, error-prone translesion replication, and nonhomologous end joining that have been implicated in interstrand DNA cross-link repair. Host cell reactivation assays monitor the overall repair capacity of a cell by providing an integrated measure of overlapping and potentially redundant processes (*21*−*23*, *27*, *47*). The efficient repair of both the CpG and GpC interstrand cross-linked plasmids clearly shows that wild-type and recombination deficient *E. coli* cell lines possess robust repair capabilities for this type of interstrand DNA cross-link. Loechler and co-workers obtained similar results using a plasmid that contained a site-specific mechlorethamine interstrand cross-link (*22*, *23*). In these studies, DNA duplexes having a single GNC step were incubated with mechlorethamine, treated with sodium hydroxide to generate the more stable formamidopyrimidine (FAPY) or imidazole ring-opened form of N^7-alkyl-guanine, and then incorporated into plasmid DNA (*22*). Repair efficiencies of these cross-linked plasmids were greater than 60% using the same *E. coli* strains used in this study and a similar repair assay. In general, the repair efficiencies in the Loechler study were somewhat lower than the repair efficiencies reported here.

NER plays an important role in the repair and tolerance of interstrand DNA cross-links (*24*, *61*). In *E. coli*, NER-mediated DNA damage recognition is an ATP-dependent process mediated by the $(UvrA)_2(UvrB)_1$ complex that recognizes damaged DNA and results in the loading of UvrB onto the DNA. Following dissociation of UvrA, the remaining UvrB recruits UvrC to the damage site which is in turn followed by DNA cleavage (*62*, *63*). The PFB50 strain of *E. coli* used in these studies contains the UvrA6 mutation (*23*), which produces a frame shift in the N-terminal part of the gene and results in a truncated UvrA protein that lacks the first ATP binding site. Consequently, the protein is not active, and thus, the entire NER pathway is nonfunctional.

The 10-fold decrease in repair efficiency observed for the CpG cross-linked plasmid DNA in the NER deficient strain is consistent with our understanding of the important role of NER in cross-link repair and is very similar to the decrease reported by Loechler and co-workers. Interestingly, when the orientation of the cross-link is reversed to give the destabilized GpC interstrand cross-link, NER-independent repair is significantly more robust. Thus, the NER-independent repair process(es) appears to be influenced by the particular structure of the cross-link-induced lesion rather than the simple presence of an interstrand cross-link.

The correlation between repair efficiency and cross-link-induced perturbations is consistent with the observation that, in general, the greater the distortion caused by a particular type of DNA damage, the higher the levels of repair. This correlation is often attributed to improved damage recognition, and this could be one of the factors contributing to the higher levels of repair observed with the GpC interstrand cross-link. While a better ability to recognize the destabilized and highly flexible GpC cross-link is one potential explanation for the increased levels of NER-independent repair, a second and perhaps complementary mechanism exists. It is possible that the stability of the cross-link itself could influence biochemical steps in the repair process. That is to say, the destabilization caused by the GpC interstrand cross-link could, in and of itself, facilitate specific mechanistic steps in the repair process such as strand separation and/or lesion bypass.

Much of what is known regarding the repair and tolerance of interstrand DNA cross-links in both bacteria and higher organisms has been ascertained by using psoralen interstrand cross-links (*14*, *15*, *25*−*27*, *64*, *65*). Psoralen is a planar aromatic photoreactive compound that preferentially intercalates and cross-links the thymines on opposite strands of -TpA- steps (*61*, *66*). Psoralen has found wide use because of its sequence preference, chemical stability, and ability to photochemically trigger the formation of cyclobutane adducts with the 5,6 double bonds of thymines (*28*). The structure of psoralen interstrand cross-links differs from that of alkyl interstrand cross-links in a number of important ways (*67*, *68*). Principally, alkyl interstrand cross-links can generally be accommodated with minor alterations to the overall helical geometry and are oftentimes associated with increases in conformational flexibility, whereas psoralen cross-links are rigidly constrained and characterized by the dramatic destabilizations associated with intercalation and helical unwinding. Differences between the processing of a psoralen interstrand cross-link and a trimethylene interstrand cross-link were observed in studies comparing the action of XPF-ERCC1 in the presence of RPA. In these studies, incision at the psoralen cross-link occurred one to five nucleotides 5′ of the cross-link, while incision at the trimethylene interstrand cross-link occurred at phosphodiester bonds five to nine nucleotides 5′ of the cross-link (*26*).

Our studies suggest that the conformational differences induced by the cross-link itself can result in considerable differences in the manner and type of repair. Indeed, it is likely that many aspects of cross-link repair, including damage recognition processes, repair pathway selection, lesion tolerance, and mutagenicity, will ultimately be influenced by the specific structure adopted by that cross-link. Studies of the type described here, which examine structure−function relationships of interstrand cross-linked DNA, may improve our understanding of the mechanisms and processes of interstrand cross-link repair and tolerance.

SUPPORTING INFORMATION AVAILABLE

Three figures showing the construction and characterization of the plasmid substrates (Figures S1−S3), one figure showing a selected spectral region from a NOESY spectrum (Figure S4), one table with the electroporation conditions used in the HCR assay (Table S1), two tables showing the NMR chemical shifts (Table S2) and geometrical parameters and backbone torsion angles for the lowest-energy calculated structure (Table S3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Erickson, L. C., Bradley, M. O., Ducore, J. M., Ewig, R. A. G., and Kohn, K. W. (1980) DNA crosslinking and cytotoxicity in normal and transformed human cells treated with antitumor nitrosoureas, *Proc. Natl. Acad. Sci. U.S.A. 77*, 467−471.
2. Garcia, S. T., McQuillan, A., and Panasci, L. (1988) Correlation between the cytotoxicity of melphalan and DNA crosslinks as detected by the ethidium bromide fluorescence assay in the F_1 variant of B_{16} melanoma cells, *Biochem. Pharmacol. 37*, 3189−3192.
3. Thomas, C. B., Osieka, R., and Kohn, K. W. (1978) DNA cross-linking by *in vivo* treatment with 1-(2-chloroethyl)-3-(4-methylcyclohexyl)-1-nitrosourea of sensitive and resistant human colon carcinoma xenografts in nude mice, *Cancer Res. 38*, 2448−2453.
4. Kano, Y., and Fujiwara, Y. (1981) Roles of DNA interstrand crosslinking and its repair in the induction of sister-chromatid exchange and a higher induction in Fanconi's anemia cells, *Mutat. Res. 81*, 365−375.
5. Colvin, M. (1982) *The Alkylating Agents*, W. B. Saunders Co., Philadelphia.
6. Kozekov, D., Nechev, L. V., Moseley, M. S., Harris, C. M., Rizzo, C. J., Stone, M. P., and Harris, T. M. (2003) DNA Interchain Cross-Links Formed by Acrolein and Crotonaldehyde, *J. Am. Chem. Soc. 125*, 50−61.
7. O'Connor, P. M., and Kohn, K. W. (1990) Comparative Pharmokinetics of DNA Lesion Formation and Removal Following Treatment of L1210 Cells with Nitrogen Mustard, *Cancer Commun. 2*, 387−394.
8. Fan, S., Chang, J. K., Smith, M. L., Duba, D., Fornace, A. J. J., and O'Connor, P. M. (1997) Cells lacking CIP1/WAF1 genes exhibit preferential sensitivity to cisplatin and nitrogen mustard, *Oncogene 14*, 2127−2136.
9. Dronkert, M. L. G., and Kanaar, R. (2001) Repair of DNA interstrand cross-links, *Mutat. Res. 486*, 217−247.
10. Ali-Osman, F., Rairkar, A., and Young, P. (1995) Formation and repair of 1,3-bis-(2-chloroethyl)-1-nitrosourea and cisplain induced total genomic DNA interstrand crosslinks in human glioma cells, *Cancer Biochem. Biophys. 14*, 231−241.
11. Dong, Q., Johnson, S. P., Colvin, O. M., Bullock, N., Kilborn, C., Runyon, G., Sullivan, D. M., Easton, J., Bigner, D. D., Nahta, R., Marks, J., Modrich, P., and Friedman, H. S. (1999) Multiple DNA repair mechanisms and alkylator resistance in the human medulloblastoma cell line D-283 Med (4-HCR), *Cancer Chemother. Pharmacol. 43*, 73−79.
12. Friedman, H. S., Colvin, O. M., Kaufmann, S. H., Ludeman, S. M., Bullock, N., Bigner, D. D., and Griffith, O. W. (1992) Cyclophosphamide resistance in medulloblastoma, *Cancer Res. 52*, 5373−5378.
13. Dong, Q., Bullock, N., Ali-Osman, F., Colvin, O. M., Bigner, D. D., and Friedman, H. S. (1996) Repair analysis of 4-hydroperoxycyclophosphamide-induced DNA interstrand crosslinking in the c-myc gene in 4-hydroperoxycyclophosphamide-sensitive and -resistant medulloblastoma cell lines, *Cancer Chemother. Pharmacol. 37*, 242−246.
14. Sinden, R. R., and Cole, R. S. (1978) Repair of cross-linked DNA and survival of *Escherichia coli* treated with psoralen and light: Effects of mutations influencing genetic recombination and DNA metabolism, *J. Bacteriol. 136*, 538−547.

15. Sinden, R. R., and Cole, R. S. (1978) Topography and kinetics of genetic recombination in *Escherichia coli* treated with psoralen and light, *Proc. Natl. Acad. Sci. U.S.A. 75*, 2373−2377.
16. Van Houten, B., Gamper, H., Holbrook, S. R., Hearst, J. E., and Sancar, A. (1986) Action mechanism of ABC excision nuclease on a DNA substrate containing a psoralen crosslink at a defined position, *Proc. Natl. Acad. Sci. U.S.A. 83*, 8077−8081.
17. Van Houten, B., Gamper, H., Hearst, J. E., and Sancar, A. (1986) Construction of DNA substrates modified with psoralen at a unique site and study of the action mechanism of ABC excinuclease on these uniformly modified substrates, *J. Biol. Chem. 261*, 14135−14141.
18. Pu, W. T., Kahn, R., Munn, M. M., and Rupp, W. D. (1989) UvrABC incision of *N*-methylmitomycin A-DNA monoadducts and cross-links, *J. Biol. Chem. 264*, 20697−20704.
19. Sladek, F. M., Munn, M. M., Rupp, W. D., and Howard-Flanders, P. (1989) In vitro repair of psoralen-DNA cross-links by RecA, UvrABC, and the 5′-exonuclease of DNA polymerase I, *J. Biol. Chem. 264*, 6755−6765.
20. Cheng, S., Van Houten, B., Gamper, H. B., Sancar, A., and Hearst, J. E. (1988) Use of psoralen-modified oligonucleotides to trap three-stranded RecA-DNA complexes and repair of these cross-linked complexes by ABC excinuclease, *J. Biol. Chem. 263*, 15110−15117.
21. Sladek, F. M., Melian, A., and Howard-Flanders, P. (1989) Incision by UvrABC excinuclease is a step in the path to mutagenesis by psoralen crosslinks in *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A. 86*, 3982−3986.
22. Berardini, M., Mackay, W., and Loechler, E. L. (1997) Evidence for a recombination-independent pathway for the repair of DNA interstrand cross-links based on a site-specific study with nitrogen mustard, *Biochemistry 36*, 3506−3513.
23. Berardini, M., Foster, P. L., and Loechler, E. L. (1999) DNA polymerase II (polB) is involved in a new DNA repair pathway for DNA interstrand cross-links in *Escherichia coli*, *J. Bacteriol. 181*, 2878−2882.
24. Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaccmaz, K., and Linn, S. (2004) Molecular Mechanisms of Mammalian DNA Repair and the DNA Damage Checkpoints, *Annu. Rev. Biochem. 73*, 39−85.
25. Bessho, T., Mu, D., and Sancar, A. (1997) Initiation of DNA interstrand cross-link repair in humans: The nucleotide excision repair system makes dual incisions 5′ to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand, *Mol. Cell. Biol. 17*, 6822−6830.
26. Mu, D., Bessho, T., Nechev, L. V., Chen, D. J., Harris, T. M., Hearst, J. E., and Sancar, A. (2000) DNA interstrand cross-links induce futile repair synthesis in mammalian cell extracts, *Mol. Cell. Biol. 20*, 2446−2454.
27. Wang, X., Peterson, C. A., Zheng, H., Nairn, R. S., Legerski, R. J., and Li, L. (2001) Involvement of nucleotide excision repair in a recombination-independent and error-prone pathway of DNA interstrand cross-link repair, *Mol. Cell. Biol. 21*, 713−720.
28. Noll, D. M., Noronha, A. M., Wilds, C. J., and Miller, P. S. (2004) Preparation of interstrand cross-linked DNA oligonucleotide duplexes, *Front. Biosci. 9*, 421−437.
29. Noronha, A. M., Wilds, C. J., and Miller, P. S. (2002) N^4C-Alkyl-N^4C cross-linked DNA: Bending deformations in duplexes that contain a -CNG- interstrand cross-link, *Biochemistry 41*, 8605−8612.
30. Noronha, A. M., Noll, D. M., and Miller, P. S. (2001) Syntheses of DNA Duplexes Containing a C-C Interstrand Cross-Link, *Nucleosides, Nucleotides Nucleic Acids 20*, 1303−1307.
31. Noronha, A. M., Noll, D. M., Wilds, C. J., and Miller, P. S. (2002) N^4C-Ethyl-N^4C Cross-linked DNA: Synthesis and Characterization of Duplexes with Interstrand Cross-links of Different Orientations, *Biochemistry 41*, 760−771.
32. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy, *J. Chem. Phys. 71*, 4546−4554.
33. Piantini, U., Sorensen, O. W., and Ernst, R. R. (1982) Multiple quantum filters for elucidating NMR coupling networks, *J. Am. Chem. Soc. 104*, 6800−6801.
34. Braunsweiler, L., and Ernst, R. R. (1987) Coherence transfer by isotropic mixing: Applications to proton correlation spectroscopy, *J. Magn. Reson. 53*, 521−528.
35. Gueron, M., and Plateau, P. (1982) Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences, *J. Am. Chem. Soc. 104*, 7310−7311.
36. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR 2*, 661−665.
37. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
38. Varani, G., Abouela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
39. Smith, J. S., and Nikonowicz, E. P. (2000) Phosphorothioate substitution can substantially alter RNA conformation, *Biochemistry 39*, 5642−5652.
40. Tjandra, N., Tate, S., Ono, A., Kainosho, M., and Bax, A. (2000) The NMR structure of a DNA dodecamer in an aqueous dilute liquid crystalline phase, *J. Am. Chem. Soc. 122*, 6190−6200.
41. Lavery, R., and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: Conventions and principles, *J. Biomol. Struct. Dyn. 6*, 655−667.
42. Brunger, A. T. (1992) *XPLOR*, Yale University Press, New Haven, CT.
43. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy minimization and dynamics calculations, *J. Comput. Chem. 4*, 187−217.
44. Clore, G. M., Gronenborn, A. M., Carlson, G., and Meyer, E. F. (1986) Stereochemistry of binding of the tetrapeptide acetyl-Pro-Ala-Pro-Tyr-NH_2 to porcine pancreatic elastase. Combined use of two-dimensional transferred nuclear Overhauser enhancement measurements, restrained molecular dynamics, X-ray crystallography and molecular modelling, *J. Mol. Biol. 190*, 259−267.
45. Berendsen, H. J. C., Postma, J. P. M., Vangusteren, W. F., Dinola, A., and Haak, J. R. (1984) Molecular-dynamics with coupling to an external bath, *J. Chem. Phys. 81*, 3684−3690.
46. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesion equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes, *J. Comput. Phys. 23*, 327−341.
47. Piette, J., Gamper, H. B., van de Vorst, A., and Hearst, J. E. (1988) Mutagenesis induced by site specifically placed 4′-hydroxymethyl-4,5′,8-trimethylpsoralen adducts, *Nucleic Acids Res. 16*, 9961−9977.
48. Wijmenga, S. S., Mooren, M. M. W., and Hilbers, C. W. (1993) in *NMR of Macromolecules, A Practical Approach* (Roberts, G. C. K., Ed.) pp 217−288, Oxford University Press, New York.
49. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
50. Hosur, R. V., Govil, G., and Miles, H. T. (1988) Application of two-dimensional NMR spectroscopy in the determination of solution conformation of nucleic acids, *Magn. Reson. Chem. 26*, 927−944.
51. Rivetti, C., Guthold, M., and Bustamante, C. (1996) Scanning force microscopy of DNA deposited onto mica: Equilibration versus kinetic trapping studied by statistical polymer chain analysis, *J. Mol. Biol. 264*, 919−932.
52. Rivetti, C., Guthold, M., and Bustamante, C. (1999) Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex, *EMBO J. 18*, 4464−4475.
53. Rivetti, C., Walker, C., and Bustamante, C. (1998) Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility, *J. Mol. Biol. 280*, 41−59.
54. Rink, S. M., and Hopkins, P. B. (1995) A mechlorethamine-induced DNA interstrand cross-link bends duplex DNA, *Biochemistry 34*, 1439−1445.
55. Webba da Silva, M., Noronha, A. M., Noll, D. M., Miller, P. S., Colvin, O. M., and Gamcsik, M. P. (2002) Solution structure of a DNA duplex containing mispair-aligned N^4C-ethyl-N^4C interstrand cross-linked cytosines, *Biochemistry 41*, 15181−15188.
56. Dooley, P. A., Zhang, M., Korbel, G. A., Nechev, L. V., Harris, C. M., Stone, M. P., and Harris, T. M. (2003) NMR determination of the conformation of a trimethylene interstrand cross-link in an oligodeoxynucleotide duplex containing a 5′-d(GpC) motif, *J. Am. Chem. Soc. 125*, 62−72.
57. Dooley, P. A., Tsarouhtsis, D., Korbel, G. A., Nechev, L. V., Shearer, J., Zegar, I. S., Harris, C. M., Stone, M. P., and Harris, T. M. (2001) Structural studies of an oligodeoxynucleotide containing a trimethylene interstrand cross-link in a 5′-(CpG) motif: Model of a malondialdehyde cross-link, *J. Am. Chem. Soc. 123*, 1730−1739.

58. Wilson, V. L., Foiles, P. G., Chung, F. L., Povey, A. C., Frank, A. A., and Harris, C. C. (2003) Detection of acrolein and crotonaldehyde DNA adducts in cultured human cells and canine peripheral blood lymphocytes by ^{32}P-postlabeling and nucleotide chromatography, *Carcinogenesis 12*, 1483−1490.
59. Edfeldt, N. B., Harwood, E. A., Sigurdsson, S. T., Hopkins, P. B., and Reid, B. R. (2004) Solution structure of a nitrous acid induced DNA interstrand cross-link, *Nucleic Acids Res. 32*, 2785−2794.
60. Edfeldt, N. B., Harwood, E. A., Sigurdsson, S. T., Hopkins, P. B., and Reid, B. R. (2004) Sequence context effect on the structure of nitrous acid induced DNA interstrand cross-links, *Nucleic Acids Res. 32*, 2795−2801.
61. Friedberg, E. C., Walker, G. C., and Seide, W. (1995) *DNA Repair and Mutagenesis*, ASM Press, Washington, DC.
62. Orren, D. K., and Sancar, A. (1989) The (A)BC excinuclease of *Escherichia coli* has only the UvrB and UvrC subunits in the incision complex, *Proc. Natl. Acad. Sci. U.S.A. 86*, 5237−5241.
63. Orren, D. K., and Sancar, A. (1990) Formation and enzymatic properties of the UvrB·DNA complex, *J. Biol. Chem. 265*, 15796−15803.
64. Li, L., Peterson, C. A., Lu, X., Wei, P., and Legerski, R. J. (1999) Interstrand cross-links induce DNA synthesis in damaged and undamaged plasmids in mammalian cell extracts, *Mol. Cell. Biol. 19*, 5619−5630.
65. Zhang, N., Lu, X., Zhang, X., Peterson, C. A., and Legerski, R. J. (2002) hMutSβ is required for the recognition and uncoupling of psoralen interstrand cross-links in vitro, *Mol. Cell. Biol. 22*, 2388−2397.
66. Pathak, M. A., Kramer, D. M., and Fitzpatrick, T. B. (1974) in *Sunlight and Man* (Kukita, A., Ed.) pp 335−368, University of Tokyo Press, Tokyo.
67. Eichman, B. F., Mooers, B. H. M., Alberti, M., Hearst, J. E., and Ho, P. S. (2001) The crystal structures of psoralen cross-linked DNAs: Drug-dependent formation of Holliday junctions, *J. Mol. Biol. 308*, 15−26.
68. Spielmann, H. P., Dwyer, T. J., Sastry, S. S., Hearst, J. E., and Wemmer, D. E. (1995) DNA Structural Reorganization Upon Conversion of a Psoralen Furan-Side Monoadduct to an Interstrand Cross-Link: Implications for DNA Repair, *Proc. Natl. Acad. Sci. U.S.A. 92*, 2345−2349.

BI050014N