Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 4580–4587
Bioorganic & Medicinal Chemistry

Intrastrand base-stacking buttresses widening of major groove in interstrand cross-linked B-DNA

Mateus Webba da Silva,[a] Ross G. Bierbryer,[a] Christopher J. Wilds,[b,†] Anne M. Noronha,[b,‡] O. Michael Colvin,[a] Paul S. Miller[b] and Michael P. Gamcsik[a,*]

[a]*Department of Medicine, Duke University Medical Center, DUMC Box 2638, Durham, NC 27710, USA*
[b]*Department of Biochemistry and Molecular Biology, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205, USA*

Received 7 February 2005; accepted 16 March 2005
Available online 9 April 2005

Abstract—The introduction of a covalent interstrand cross-link induces changes in the intrinsic structure and deformability of the DNA helix that are recognized by elements of the DNA repair apparatus. In this context, the solution structure of the undecamer d(CGAAAT*TTTCG)$_2$, where T* represents a N^{3}T–butyl–N^{3}T interstrand cross-link, was determined using molecular dynamics calculations restrained by NOE and dihedral angle data obtained from NMR spectroscopy. The structure of this cross-linked undecamer shows dramatic widening of the major groove of the B-DNA stem without disruption of Watson–Crick base pairing. This change in tertiary structure illustrates the cumulative effect of cooperativity in intrastrand base stacking of an A-tract of three adenines. Further, it is the direct result from the imposition of geometric angular constraints by the cross-link chain on an ApT* and T*pT steps in the segment AAAT*T. The widening of the major groove is due to the dominant contribution of base stacking to the stability of the ApT compared to the TpT step suggesting that the latter is more deformable within a DNA stem. Compared to earlier structures of ethyl cross-linked oligonucleotides, this unique perturbation induced by the butyl moiety offers a new probe for systematic studies of DNA repair mechanisms.

1. Introduction

Human cells are constantly assaulted by both endogenous and exogenous agents that can link opposite strands of DNA resulting in lesions that cause dramatic consequences to a cell if left unrepaired.[1] The cytotoxicity of bifunctional alkylating anticancer agents is likely due to the formation of such linkages that inhibit transcription and replication.[2] Today, cross-linking agents such as the nitrogen mustards (melphalan, chlorambucil, cyclophosphamide, and ifosfamide), the alkylalkanesulfonate busulfan, chloroethylnitrosoureas (carmustine and lomustine), the platinum drugs (cisplatin and carboplatin), and the natural product mitomycin C are among the most widely used and most effective anticancer agents, as both single agents or as part of combination therapy regimens. For these drugs the interstrand cross-link is believed to be the critical cytotoxic lesion. Increased repair of these lesions has been observed in drug-resistant cancer cells.[3] Circumvention of this resistance will require an understanding of the factors involved in its recognition and repair. Therefore, the molecular mechanisms involved and the structural determinants of cellular sensitivity to cross-linking agents are of fundamental biological importance.

Our efforts are focused upon how repair proteins distinguish interstrand cross-link lesions from the vast excess of normal DNA. The results of these studies should allow us to structurally define repair susceptibility for interstrand cross-links leading to more effective anticancer agents that are resistant to repair. Our main premise is that both local structure and sequence dependent

Abbreviations: COSY, 2D correlated spectroscopy; C*, N^{4}C–ethyl–N^{4}C; DQF-COSY, double quantum filtered COSY; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; rMD, restrained molecular dynamics; TOCSY, 2D total correlated spectroscopy; T*, N^{3}T–butyl–N^{3}T

Keywords: DNA Interstrand crosslinks; DNA repair; Solution NMR structure.

* Corresponding author. Tel.: +1 (919) 681 2244; fax: +1 (919) 668 3925; e-mail: michael.gamcsik@duke.edu
† Present address: Department of Chemistry and Biochemistry, Concordia University, Montreal, Quebec, Canada.
‡ Present address: Alnylam Pharmaceuticals, Cambridge, MA, USA.

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.03.032

deformability of DNA segments can enhance or reduce their ability to fold into nucleosomes, or to be recognized by proteins.[4] In order to understand what structural properties will lead to the assembly of specific repair apparatus and their specificities, we are systematically creating a number of well characterized probes bearing specific structural and biophysical properties.[5–10] Essential to our effort has been the development of synthetic strategies to prepare cross-linked oligonucleotides with defined structure.[7,10] In this regard, we have previously reported the synthesis and physical characterization of a series of short DNA duplexes that contain N^3T–alkyl–N^3T interstrand cross-links (alkyl = ethyl, butyl, heptyl).[10] Detailed structures of oligomers containing ethyl cross-links derived from NMR data have been presented.[9,10] This report details the unique structural perturbations caused by increasing the cross-link chain length.

The undecamer containing a mispair aligned N^3T–butyl–N^3T used in this study is shown in Figure 1a. This differs only in cross-link chain length from our previous report of this undecamer containing a mispair aligned N^3T–ethyl–N^3T moiety (Fig. 1b).[8] Although proceeding from an ethyl to butyl cross-link appears to be only small change in covalent structure, there are dramatic changes in tertiary conformation. Further comparisons of the structure of this oligomer with that of mispair aligned N^4C–ethyl–N^4C (hereafter denoted C*) (Fig. 1c)[9] in the same scaffold, 5′-CGAAAPy*TTTCG-3′ show that each of these relatively minor changes in chain length and identity of the pyrimidine base pair results in unique structural changes that offer distinct targets for recognition and repair.

2. Materials and methods

The oligomer d(CGAAAT*TTTCG)$_2$ was synthesized and purified using methods previously described to prepare short DNA duplexes that contain an N^3T–alkyl–N^3T interstrand cross-link.[10] The duplex gave the expected nucleoside to cross-link ratios when digested with a combination of snake venom phosphodiesterase and calf intestinal alkaline phosphatase, followed by analysis by reversed phase HPLC. The mass of the duplex was consistent with its structure as analyzed by mass spectrometry (*m/z*: calcd 6717.08, found 6718.78). The NMR sample was prepared by dissolving the molecule in 0.25 mL buffer containing 15 mM sodium phosphate (pH 7.4) and 100 mM NaCl resulting in a solution approximately 3 mM in single strand concentration.

2.1. NMR Spectroscopy

Standard NMR experiments were recorded on Varian Inova 500, 600, and 800 MHz spectrometers. We recorded NOESY[1] (200 and 60 ms) spectra in ^1H_2O at 0 °C and NOESY[11] (50, 100, 150, 200, and 250 ms), DQF-COSY.[12] TOCSY[13] (spin-lock time of 40 ms) in ^2H_2O at 20 °C. In ^1H_2O, data were acquired with a jump-and-return pulse sequence[14] and in ^2H_2O, with Watergate suppression of the residual water signal.[15] All data sets were acquired in a phase sensitive mode (TPPI). In ^1H_2O, NOESY data sets were collected with 8*k* 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 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 from Insight II (Accelrys Inc.).

2.2. Distance and torsion restraints

NOE cross-peaks involving exchangeable protons in NOESY spectra (50 and 200 ms mixing times) in H_2O

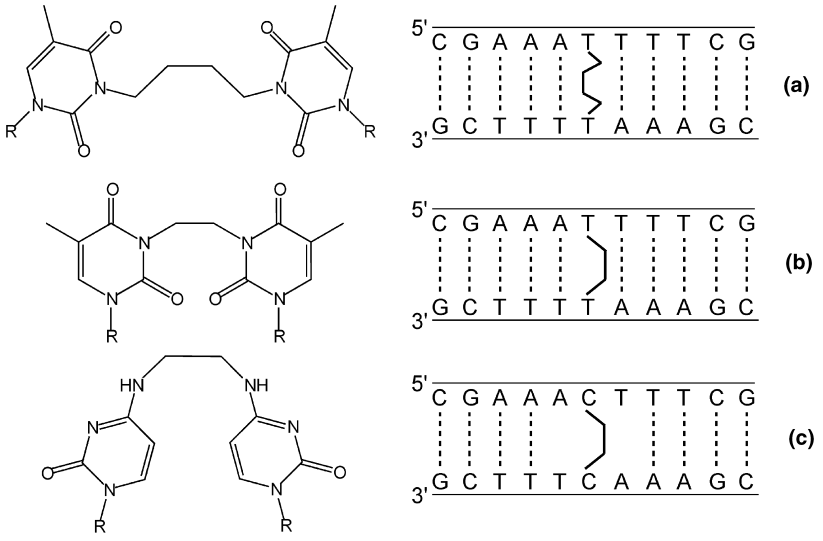

Figure 1. Scheme of self-complementary duplexes within the scaffold 5′-CGAAAPy*TTTCG-3′ with corresponding chemical structure for (a) Py* = N^3T–butyl–N^3T, (b) Py* = N^3T–ethyl–N^3T, and (c) Py* = N^4C–ethyl–N^4C.

buffer were classified as strong (strong intensity at 60 ms), medium (barely observable at 60 ms) and weak (not observable at 60 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 non-exchangeable protons were derived from NOESY spectra in ^2H_2O buffer recorded as a function of mixing time (50, 100, 150, 200, 250 ms). Distances were estimated from the initial build-up rates within FELIX 2000 (Accelrys). The cytosine H5:H6 interproton distance of 2.46 Å was used as 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[16] 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. Thus 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*. The program CURVES 5.2[17] was used to estimate DNA conformation and helical parameters.

2.3. Distance restrained molecular dynamics regularization

Calculations were performed with XPLOR version 3.1[18] and XPLOR-NIH[19] using the CHARMm force field[20] and adapted for restrained molecular dynamics (rMD) for nucleic acids. All calculations were executed in vacuo without explicit counter ions. The initial distance geometry and simulated annealing refinement protocol started from 200 different structures generated from sets of two strands, each 11 nucleotides long, randomized over all dihedral angles. A number of structures (23 out of 200) emerged separated from non-converged structures by gaps in components of the potential energy function (dihedral angles, van der Waals, NOE violations, and covalent geometry). This set was subsequently submitted to sets of rMD calculations performed using random velocities fitting a Maxwell–Boltzmann distribution as described previously.[8,9,21] Coordinates (accession number: 1XCI) have been deposited in the PDB data bank.

3. Results

Well-resolved sharp peaks for three lower field thymine imino, two guanine imino, and two amino peaks are observable in the region between 8 ppm and 14.5 ppm indicating a stable, one conformer structure suitable for NMR studies (Fig. 2a). The C_2 symmetry of this molecule causes equivalent positions to exhibit identical chemical shifts in the spectra. Thus complementary base pairing is assumed throughout the stem with an exception for the symmetrical T6*–T6* mispair.

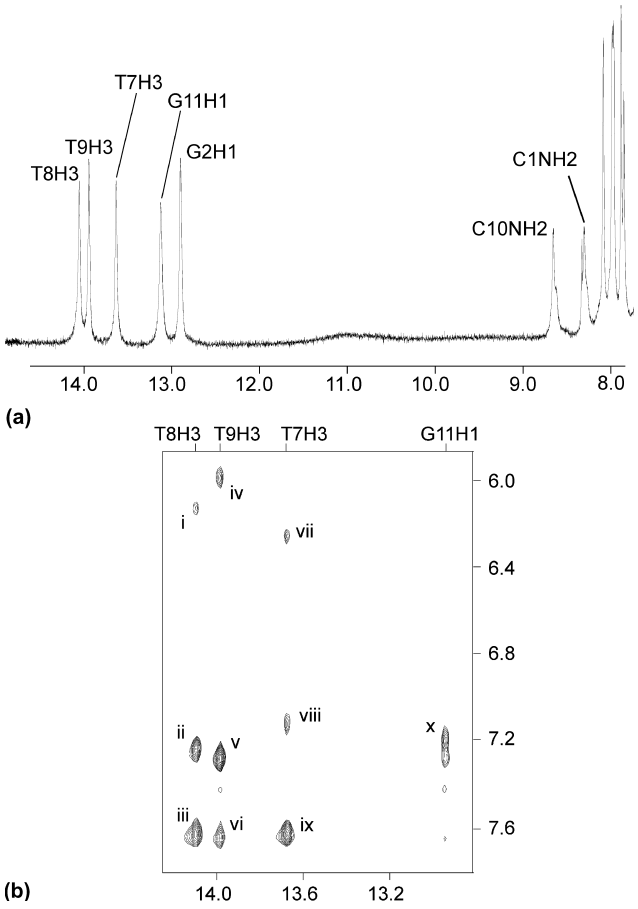

Figure 2. (a) One dimensional ^1H NMR spectrum of d(CGAAAT*TTTCG) T* = N^3T–butyl–N^3T in ^1H_2O solution containing 0.1 M NaCl at pH 7.4 showing assigned imino proton signals. (b) Selected spectral regions from a NOESY (200 ms) in ^1H_2O at 0 °C, pH 7.4, and 0.1 M NaCl for d(CGAAAT*TTTCG) T* = N^3T–butyl–N^3T. The assignments indicated are typical for establishing canonical base pairing: (i) T8H3–AH62, (ii) T8H3–A4H2, (iii) T8H3–A4H61, (iv) T9H3–A3H62, (v) T9H3–A3H2, (vi) T9H3–A3H61, (vii) T7H3–A5H62, (viii) T7H3–A5H61, (ix) T7H3–A5H2, (x) G11H1–C10H42.

3.1. Resonance assignments

Proton signal assignments were made through analysis of standard NOESY, TOCSY, and DQF-COSY spectra by established methods.[22,23] Assignment of the imino and amino exchangeable protons followed analysis of 1–1 NOESY spectra at 60 and 200 ms mixing time and at 0 °C, and is based on assignments of the non-exchangeable region. Figure 2b shows a selected region of a 1–1 NOESY spectrum (200 ms) depicting assignments for selected exchangeable protons. Typical cross-strand cross-peaks defining G:C and A:T base pairs were observed and appear in Figure 2b: A3 (T9H3:A3H61, T9H3:A3H62, T9H3:A3H2), A4 (T8H3:A4H61, T8H3:A4H62, T8H3:A4H2), A5 (T7H3:A5H61, T7H3:A5H62, T7H3:A5H2), C1 (G11H1:C1H42, G11H1:C1H41), and C10 (G2H1:C10H42, G2H1:C10H41), T8H3–A5H2, T9H3–A4H2. The DNA duplex is right handed with all its bases in an *anti* glycosidic bond angle as established by the ready observation of sequential connectivities between the base and

its own 5′-flanking sugar H1′ protons along individual strands (Fig. 3a), as well as by observation of H6/H8-H3′, by aromatic–aromatic cross-peaks, and H6/H8-H2′/H2″ connectivities. Typical B-DNA cross-strand (i)AH2:(j + 1)NH1′ (i and j represent different strands, N is any base; A4H2:T9H1′, A5H2:T8H1′, A3H2:C10H1′), intrastrand (i)AH2:(i+1)NH1′ (A4H2:A5H1′, A5H2:T6H1′, A3H2:A3H1′), and intra-residue (A4H2:A4H1′, A5H2:A5H1′, A3H2:A3H1′) are also observable (Fig. 3).

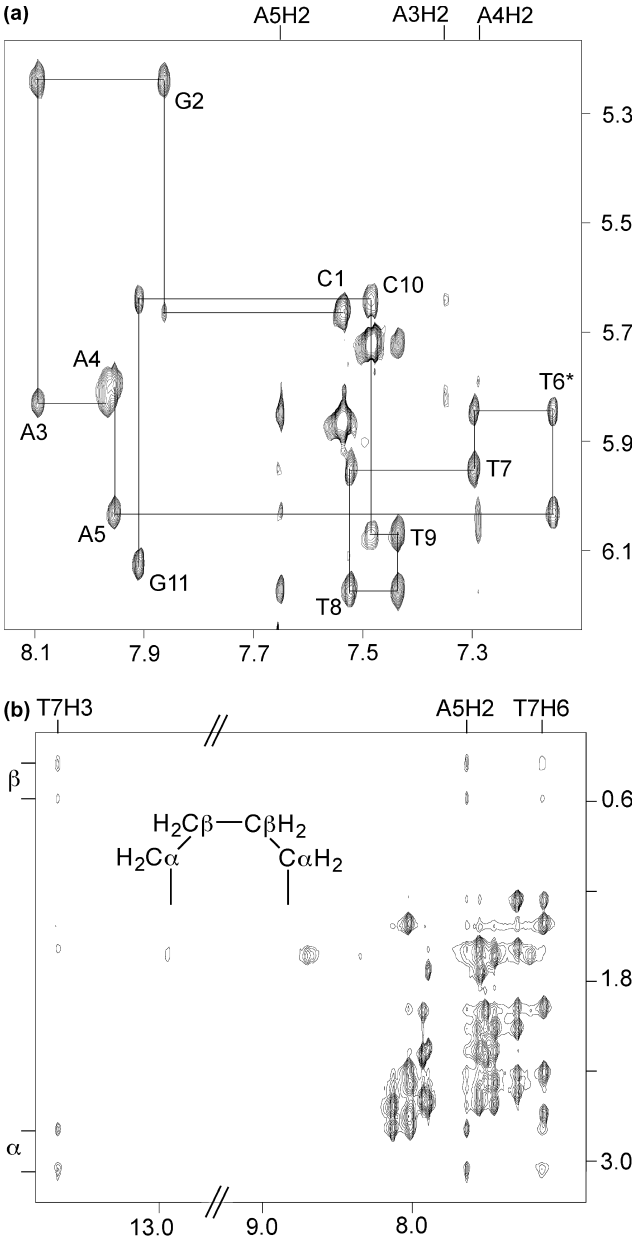

Figure 3. (a) Selected region of NOESY (250 ms mixing time) spectrum d(CGAAAT*TTTCG) T* = N^3T–butyl–N^3T in ^2H_2O solution at pH 7.4, 0.1 M NaCl 20 °C showing NOE correlations between H8/H6/H2/H5 and H1′. The self-peaks involving H1′ protons have been labeled and their sequential connectivities have been drawn in. (b) Selected region of NOESY (250 ms mixing time) spectrum d(CGAAAT*TTTCG) T* = N^3T–butyl–N^3T in ^2H_2O solution at pH 7.4, 0.1 M NaCl 20 °C depicting NOE correlations involving the exocyclic cross-linking butyl moiety labeled as shown in the inset.

Sugar geometries populate predominantly S-type[24] as derived from stereospecific assignment of individual H2′ and H2″ protons initially derived from comparison of intensity of the H1′:H2′ and H1′:H2″ intraresidue cross-peaks in a 50 ms NOESY spectrum. This was confirmed from DQF-COSY spectral patterns in which the H1′:H2″ coupling constants were reasonably large (>6 Hz) with only one terminal residue (G11) showing strong H2″:H3′ cross-peaks. The latter was left unrestrained during structure calculations.

The environment of the alkylated base within the DNA stem is well characterized through a good number of dipolar connectivities. Figure 2b depicts some of the connectivities of the cross-linking butyl moiety. Cross-peaks of the butyl moiety to T7H3, together with the lack of cross-peaks to any sugar pucker or backbone protons indicate that it is not minor groove orientated (Fig. 3b). Noticeably stronger cross-peaks are observed to the α-methylene protons as compared to the β-methylene. This could reflect cross-strand interactions since the T7:A5 base pair is unperturbed. Indeed, along with this assertion is the fact that cross-peaks A5H2–T6*CαH_2 are perceptibly stronger than A5H2–T6*CβH_2. Notably, cross-peaks of the butyl moiety to T7H6 are weaker than to A5H2 (Fig. 3b).

3.2. Structural features

A stereoview of 15 superpositioned refined structures of the duplex d(CGAAAT*TTTCG)$_2$ is plotted in Figure 4 and exhibits average heavy atom root-mean-square-deviations (rmsd) to the mean structure of 0.7 ± 0.2 Å. The input and structure convergence parameters are listed in Table 1. The right-handed DNA shows little departure from regularity along the backbone (Fig. 5a

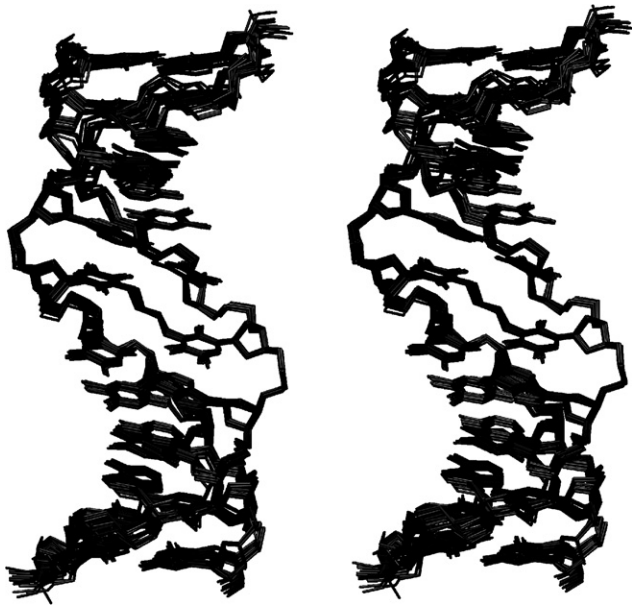

Figure 4. Stereoview of a superposition of 15 lowest energy refined structures of the duplex d(CGAAAT*TTTCG)$_2$ T* = N^3T–butyl–N^3T. For clarity protons and phosphorus-attached oxygen atoms have been removed.

Table 1. Restraints and refinement statistics for 15 selected structures for d(CGAAAT*TTTCG)$_2$

A. NMR restraints	
Total number of restraints	283
Non-exchangeable protons	155
Exchangeable protons	36
Hydrogen bond restraints (empirical)	48
Dihedral angle restraints (pucker only)	44
Non-crystallographic symmetry restraints on all heavy atoms	
B. Structural statistics	
NMR *R*-factor ($R_{1/6}$):	0.106–0.107
NOE rmsd (Å) total	0.09–0.090
NOE violations exceeding 0.2 Å	0
Heavy-atoms pairwise rmsd	0.7 ± 0.2

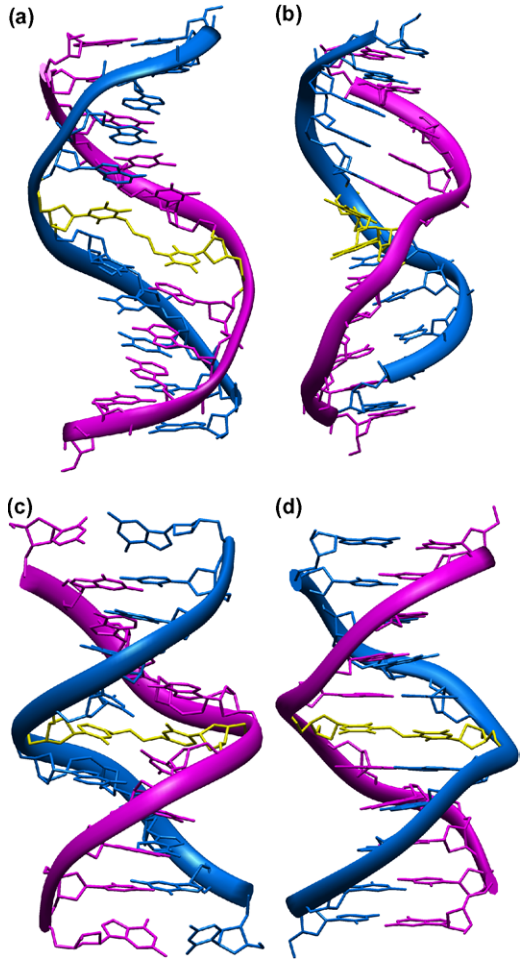

Figure 5. Views depicting the length of the stem and the regularity of the backbone. The individual strands are colored pink and blue and the cross-link is colored yellow. (a) and (b) orthogonal views of the duplex structure of d(CGAAAT*TTTCG)$_2$ T* = N^3T–butyl–N^3T. Major groove view of the structure of d(CGAAAPy*TTTCG)$_2$ (c) Py* = N^4C–ethyl–N^4C, and (d) Py* = N^3T–ethyl–N^3T.

and b). The end-to-end pitch is typical of B-DNA (~35 vs 34 Å), and a stem mean magnitude for intrastrand inter-base twist of 39°; that is, overwound by about 3° as compared to B-DNA (~36°).[16] Without disruption of base pairing throughout the stem, the cross-link site bears remarkable influence on the tertiary structure by dramatically widening the major groove and inducing a small global axis curvature of 7° (Fig. 5a and b).

In contrast to all other bases there is a perceptible measure in the magnitude of base-pair stagger (2°), buckle (3°), and large propeller twisting (35°) for T6*–T6* mispair; Figure 6b. These deviations are coupled with larger magnitudes for the T6*pT7 base step parameters such as rise, tilt, and twist (Table 2). The larger magnitude for stepwise twisting for T6*pT7 is coupled to a noticeably smaller magnitude for the preceding base step, A5pT6*. These two latter facts are at the basis of the small global axis curvature observed (Fig. 5b). These local distortions reflect the disruption of intrastrand base stacking in the step T6*pT7 whilst maintaining good intrastrand stacking between T6* and A5 (Fig. 6a and b). Excellent stacking between A5 and T6* is not surprising since purine-to-pyrimidine steps are expected to exhibit extensive intrastrand overlap.

In summary the very localized distortions cause a wider major groove and some axis curvature, coupled with sugar puckers populating predominantly S-type conformations, and thus lead to an overall canonical B-DNA topology. No disruptions were observed to the canonical base pairing and a small measure of stem curvature is observed. Local environments, as determined from sugar puckering and other geometrical parameters, remain typical of B-DNA. Distortions were observed on the base step stagger, shear and opening specifically associated with the tethered base. Due to the short distance between the N3 atoms of the thymines, geometric constraints due to the sp^3-hybridized carbons of the exocyclic butyl induce departure from co-planarity for these mismatched bases, therefore disrupting base juxtaposition for the base step T6*pT7. However all other, subsequent and preceding, intrastrand base stacking stabilizing interactions remain; including the base step ApT6*. The increase in magnitudes of the T6*pT7 base step rise and tilt, and the concomitant small A5pT6* base step twist, reflect a wider major groove. The major groove widening is well characterized. Typical interstrand cross-peaks indicative of B-DNA base-step as observed from the major groove for the three-adenine segment were not observed; that is, cross-peaks for T8H3–A5H2, T9H3–A4H2 are missing in the 200 ms mixing time spectrum. However, the typical cross-strand A4H2:T9H1′, A5H2:T8H1′, and A3H2:C10H1′ were observed. Together these facts clearly establish major groove widening. Notably, although the interstrand cross-peaks (T8H3–A5H2 and T9H3–A4H2) were not observed in the 200 ms NOESY experiment, we did not impose repulsive restraints at any stage of the structure calculations.

4. Discussion

The double helical structure of DNA is strongly favored enthalpically and base stacking interactions are believed to be the dominant contributor to the stabilization of

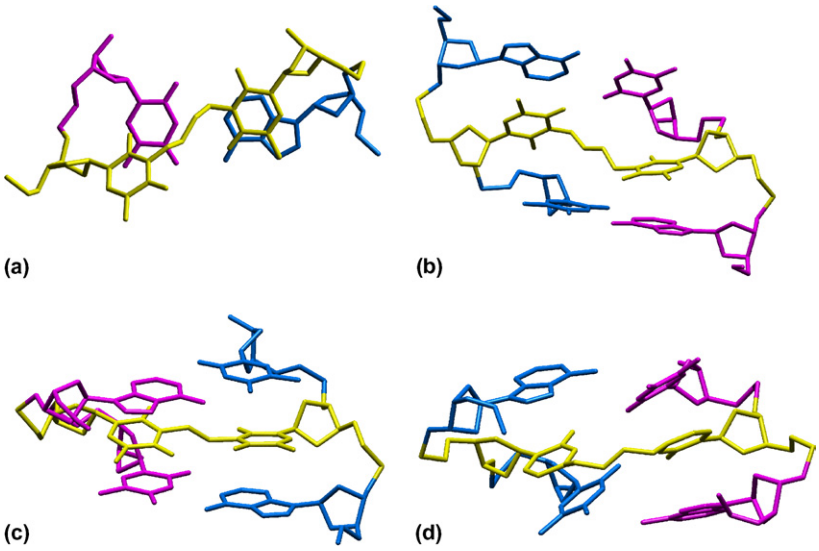

Figure 6. (a) View of stacking juxtaposition of a A5pT6* base step in the structure of d(CGAAAT*TTTCG)$_2$ T* = N^{3}T–butyl–N^{3}T. Views into the stem axis showing the alignment of the mispair aligned bases with the (b) butyl and (c) ethyl moieties in d(CGAAAT*TTTCG)$_2$, and (d) ethyl cross-link in d(CGAAAC*TTTCG)$_2$.

Table 2. Geometrical base step parameters for the lowest energy structure of d(CGAAAT*TTTCG)

Base step	Rise (Å)	Tilt[a] (Å)	Twist (Å)
C1pG2	2.9	2.2	46.6
G2pA3	3.2	0.2	46.8
A3pA4	3.3	2.2	32.4
A4pA5	3.2	1.7	34.0
A5pT6*	3.1	4.8	11.6
T6*pT7	5.2	18	48.7
T7pT8	3.3	0.3	36.2
T8pT9	3.5	3.6	33.1
T9pC10	3.2	0.2	47.6
C10pG11	3	1.8	46.1

[a] Absolute magnitude.

the helix.[25] The stability imparted by base stacking is related to the size of the flat π-surface of the stacked aromatic moieties. The increased stabilization of the double helix imparted by base stacking offsets the less favorable backbone conformations that are necessary to orient the stacked bases. Usually conformational coupling of sugar phosphate backbone and base stacking interactions determine deformability of DNA. In the present case the entropically favored sugar phosphate backbone is virtually unperturbed by the lesion maintaining its normal B-DNA characteristics. We have here established that interstrand cross-links can induce local distortions that dominated by base stacking interactions propagate to determine global helical conformations without deviation from B-DNA. Indeed, geometrical constraints of the cross-link chain force a large propeller twist on the cross-linked mispair. A large T*/T* base mispair propeller twisting (49°) forces a kink within the DNA toroid resulting in an abrupt opening of the major groove centered at the T*pT step of the mispair (Fig. 5b). Thus the cross-link induces a major perturbation at the T*pT step since changes in translation parameters are strongly penalized, especially those that lead to loss of stacking.[26] Notably, the widening effect is not at the expense of backbone compression (minor groove narrowing). In contrast, the presence of AAAT* segments on both sides of the lesion in the pallindromic structure allows for cooperativity effects to add constructively, thus resulting in an overall deflection of the helix axis with a starting point at the lesion site. The dominant intrastrand base stacking of ApT* is coupled to a three residue A-tract that stabilizes intrastrand stem progression away from the lesion and results in a widening of the major groove within the central seven base pairs of the DNA stem. The strong compression of this short A-tract toward the 3′-end, coupled with dominant base-stacking in the ApT* compared to T*pT step, results in a wider major groove in the helical global topology. Indeed, DNA base-pair deformability studies inferred from molecular dynamics simulations have shown that in the order of deformability of dinucleotide steps, pyrimidine–purine steps are the most deformable and ApT is the stiffest of the dinucleotide steps.[27] An A-tract containing an ApT step at its 3′ end is unusually rigid because the thymine methyl group clashes sterically with the sugar/phosphate backbone, blocking roll compression toward the major groove.[28]

Our results are further proof of the fragility of TpT sites in DNA structure. These steps are usually natural fracture points for DNA–bending proteins.[29] By systematically changing the nature of tethered residue and/or length of cross-link chain in the same scaffold, we probed DNA stem plasticity. In comparing the structure of the mispair aligned N^{3}T–butyl–N^{3}T and N^{3}T–ethyl–N^{3}T[8] interstrand cross-links it is striking that whilst in the former there is a major groove widening, there is no perceptible change in the dimensions of the grooves in the latter (Fig. 5a and d). Furthermore, the ethyl tether does not induce axis curvature in the stem. It is situated *between* minor and major grooves in a manner perpendicular to the stem axis (Fig. 5c and d). In

contrast to the ethyl cross-link, the configurational constraints of the butyl tether perturb the structure, also resulting in accommodation in a region between major and minor grooves, but in a staggered fashion; that is, deviating from perpendicularity to the stem axis (Figs. 5a and 6b). For both, N^3T–ethyl–N^3T and N^4C–ethyl–N^4C[9] interstrand cross-links, the magnitude of base pair buckle and propeller twist diminishes progressively away from the lesion site. However for the N^3T–butyl–N^3T cross-link, the base pair buckle and twist parameter perturbations are localized only to the tethered and adjacent bases. Thus, compared to the ethyl cross-link, the butyl linker destabilizes the DNA scaffold only by about 6 °C[10] (i.e., its melting temperature is about 6 °C higher). This is probably illustrative of the fact that the stacking energetics was not significantly altered.

Notably, both N^3T–butyl–N^3T and N^4C–ethyl–N^4C interstrand cross-links induce changes in the stem structure albeit rather differently (Fig. 5a–c). The major groove protruding ethyl moiety in N^4C–ethyl–N^4C pulls the strands apart in a manner perpendicular to the stem axis, thus collapsing the stem into a 27° bend. However, the butyl moiety in N^3T–butyl–N^3T induces an intrastrand staggering collapse leading to a butyl moiety not perpendicular to the stem axis and with concomitant stem curvature of 7° toward the minor groove. The differences in magnitude of stem axis bending are further modulated by rupture of contiguous base pairs to the N^4C–ethyl–N^4C and no disruptions in hydrogen bond alignments for N^3T–butyl–N^3T (Figs. 1a and c 6a and d). Of interest is also the fact that by pulling the strands apart the lesion disruption for the N^4C–ethyl–N^4C interstrand cross-link induces *x*-displacement with concomitant population of C1′-*endo* sugar puckers. This is a clear manifestation of B- to A-DNA conversion. This contrasts to N^3T–butyl–N^3T in which we observe canonical B-DNA local structure through out the stem.

Perturbation of intrinsic structure of DNA as well as its deformability shape the nature and efficiency of molecular mechanisms of DNA repair. Proteins that recognize DNA damage carry out their function in the presence of a large conformational space of structural variation. In these conditions both local and long-range distortions to DNA conformation are expected to affect normal protein–DNA interactions. Since many proteins bind DNA through the major groove, changes in its dimensions provide an important means of regulating accessibility of proteins. Our current structure thus provides a means for probing the effect of major groove widening on the event of recognition. Bifunctional alkylating agents that induce this type of distortion without change in local DNA structure may likely convene a specific repair apparatus, and therefore might evade repair mechanisms based on local DNA distortions. Compared to our previous work, the current study provides a unique specific structural perturbation that should be determinant for differences in both nature and efficiency of DNA repair apparatus. Progress in elucidating the mechanisms of DNA repair pathways involved in maintaining genome stability can benefit from these structural approaches.

Acknowledgements

This research was supported by a grant from the National Cancer Institute CA082785 (to P.S.M.) and a postdoctoral fellowship from the Natural Sciences and Engineering Council (to A.M.N.). The Duke University NMR Center was established with grants from the National Institutes of Health, National Science Foundation and the North Carolina Biotechnology Center.

Supplementary data

A table with geometrical base pair parameters for the lowest energy calculated structure as well as one figure showing a DQF-COSY spectrum of d(CGA-AAT*TTTCG)$_2$ at 20 °C. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2005.03.032.

References and notes

1. Dronkert, M. L.; Kanaar, R. *Mutat. Res.* **2001**, *486*, 217–247.
2. Colvin, M.; Chabner, B. A. In *Cancer chemotherapy: Principles and Practice*; Collins, J. M., Ed.; J.B. Lippincott: Philadelphia, 1990, pp 276–313.
3. 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.; Friedman, H. S. *Cancer Chemother. Pharmacol.* **1999**, *43*, 73–79.
4. Olson, W. K.; Gorin, A. A.; Lu, X.-J.; Hock, L. M.; Zhurkin, V. B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11163–11168.
5. Noll, D. M.; Noronha, A. M.; Miller, P. S. *J. Am. Chem. Soc.* **2001**, *123*, 3405–3411.
6. Noronha, A. M.; Wilds, C. J.; Miller, P. S. *Biochemistry* **2002**, *41*, 8605–8612.
7. Noronha, A. M.; Noll, D. M.; Wilds, C. J.; Miller, P. S. *Biochemistry* **2002**, *41*, 760–771.
8. Webba da Silva, M.; Wilds, C. J.; Noronha, A. M.; Colvin, O. M.; Miller, P. S.; Gamcsik, M. P. *Biochemistry* **2004**, *43*, 12549–12554.
9. Webba da Silva, M.; Noronha, A. M.; Noll, D. M.; Miller, P. S.; Colvin, O. M.; Gamcsik, M. P. *Biochemistry* **2002**, *41*, 15181–15188.
10. Wilds, C. J.; Noronha, A. M.; Robidoux, S.; Miller, P. S. *J. Am. Chem. Soc.* **2004**, *126*, 9257–9265.
11. Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4554.
12. Piantini, U.; Sorensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 6800–6801.
13. Braunsweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1987**, *53*, 521–528.
14. Gueron, M.; Plateau, P. *J. Am. Chem. Soc.* **1982**, *104*, 7310–7311.
15. Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661–665.
16. Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984.
17. Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 655–667.

18. Brunger, A. T. version 3.1. A system for X-ray crystallography and NMR. Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, 1992.
19. Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Res.* **2003**, *160*, 66–74.
20. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
21. Webba da Silva, M. *Biochemistry* **2003**, *42*, 14356–14365.
22. Wijmenga, S. S.; Mooren, M. M. W.; Hilbers, C. W. In *NMR of Macromolecules, A Practical Approach*; Roberts, G. C. K., Ed.; Oxford University Press: New York, 1993, pp 217–288.
23. Wuthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
24. Hosur, R. V.; Govil, G.; Miles, H. T. *Magn. Reson. Chem.* **1988**, *26*, 927–944.
25. Opresko, P. L.; Sweasy, J. B.; Eckert, K. A. *Biochemistry* **1998**, *37*, 2111–2119.
26. Perez, A.; Noy, A.; Lankas, F.; Luque, F. J.; Orozco, M. *Nucl. Acids Res.* **2004**, *32*, 6144–6151.
27. Lankas, F.; Sponer, J.; Langowski, J.; Cheatham, T. E. *Biophys. J.* **2003**, *85*, 2872–2883.
28. Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025–1054.
29. Luger, K.; Mader, A. W.; Richmond, R. K.; Sargent, R. G.; Richmond, T. J. *Nature* **1997**, *389*, 251–260.