A structural snapshot of base-pair opening in DNA

Daan M. F. van Aalten*, Daniel A. Erlanson†‡, Gregory L. Verdine†, and Leemor Joshua-Tor*§

*W. M. Keck Structural Biology Laboratory, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724; and †Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138

Communicated by James D. Watson, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, July 30, 1999 (received for review May 21, 1999)

The response of double-helical DNA to torsional stress may be a driving force for many processes acting on DNA. The 1.55-Å crystal structure of a duplex DNA oligonucleotide d(CCAGGCCTGG)$_2$ with an engineered crosslink in the minor groove between the central guanine bases depicts how the duplex can accommodate such torsional stress. We have captured in the same crystal two rather different conformational states. One duplex contains a strained crosslink that is stabilized by calcium ion binding in the major groove, directly opposite the crosslink. For the other duplex, the strain in the crosslink is relieved through partial rupture of a base pair and partial extrusion of a cytosine accompanied by helix bending. The sequence used is the target sequence for the *Hae*III methylase, and this partially flipped cytosine is the same nucleotide targeted for extrusion by the enzyme. Molecular dynamics simulations of these structures show an increased mobility for the partially flipped-out cytosine.

The cellular genome is subjected to strain forces that can drive the DNA double helix to adopt a variety of unusual structures. Many essential operations on the genome such as transcription, replication, recombination, repair, and methylation proceed through transient intermediates in which DNA structure deviates sometimes radically from the canonical *B* form duplex. Because the cellular context often provides the conditions necessary to drive the formation of such torsionally stressed structures, or because of their fleeting nature, few have been captured in a form that has enabled high-resolution structural studies. As a consequence, little detailed information is available regarding how naked DNA responds to torsional stress.

Perhaps the most visually striking distortion brought on DNA by the binding of a protein is so-called "base flipping," wherein an entire nucleoside unit is swiveled out of the DNA helix and inserted into a recognition pocket on the protein. Extrahelical, or flipped-out, bases were initially observed in duplex DNA in a crystal structure (1) and several NMR studies of *B* form DNA containing unpaired bases (2–4). The first flipped bases in protein complexes were observed for the substrate cytosine residue of the DNA (cytosine-5)-methyltransferases (m5C-MTases) M.*Hha*I (5) and M.*Hae*III (6). Nucleoside extrusion and extrahelical recognition have since been directly implicated in the catalytic mechanism of diverse DNA-processing enzymes. These are exemplified by crystal structures of 3-methyladenine DNA glycosylase (Aag) (7), T4 endonuclease V (8), uracil-DNA glycosylases (9, 10), and G:T/U mismatch-specific DNA glycosylase (11), all complexed to DNA, and are suggested to occur in numerous other cases (12). These findings have aroused interest in understanding processes through which torsional stress imposed on DNA can induce the opening of base pairs.

Verdine and coworkers have developed a strategy for trapping torsionally stressed structures in oligonucleotides. In this system, disulfide crosslinks are engineered site-specifically into DNA in such a way as to pose a particular conformational challenge, to which the double helix responds by adopting the most energetically accessible noncanonical structure. The thiol-bearing tethers used to form crosslinks are attached at positions that do not themselves alter duplex DNA structure, thus ensuring that the final structures primarily reflect a response to torsional stress induced by crosslinking. Disulfide crosslinking has been used, for example, to enforce the formation of a localized 30° bend in duplex DNA, thereby generating a native-like binding motif for the architecture-specific chromatin remodeling protein HMG-D (13). Before structures of m5C-MTase·DNA complexes were available, interstrand disulfide crosslinking was used to probe the interaction of the enzyme with its target recognition sequence (14, 15). Disulfide-bonded alkyl tethers were attached to the central guanines (**G**), such as in Fig. 1 *A* and *B*, of the palindromic *Hae*III site (5′-G**G***C*C-3′·3′-C*C***G**G), which are directly base paired to the substrate cytosine residues (*C*); by virtue of tether attachment at the N2 atom of **G**, the crosslink was spatially constrained to the minor groove. Although this modification *per se* had little effect on the rate of enzymatic methylation (14), it exerted a profound and specific influence on binding by the enzyme. Namely, crosslinking via an ethyldisulfide (C_2) tether increased the strength of M.*Hae*III-DNA interactions by ≈50-fold, whereas a propyldisulfide (C_3) crosslink did not stimulate binding by the enzyme (14, 15). To explain these results, it was proposed that the C_2 crosslink was too short to span its attachment points in a relaxed DNA duplex and therefore could be accommodated only through induction of a torsionally stressed structure in DNA, which in turn reduced the energetic cost of further DNA deformation by the enzyme. Elongation of the tether by two methylene units (C_3 crosslink) was thus suggested to relieve crosslink-induced helical stress. Consistent with this interpretation, the NMR spectra of the C_3 crosslink exhibited features characteristic of normal base pairing, whereas those of the C_2 crosslink were suggestive of lost or aberrant pairing (15, 16).

Here we describe the crystal structure of the C_2 crosslinked decamer, d(CCAG**G**CCTGG)$_2$, where **G** denotes the C_2 crosslinked guanines (Fig. 1 *A* and *B*), at 1.55 Å resolution. The decamer crystallizes with two distinct molecules in the asymmetric unit, one having an essentially canonical duplex and apparently strained crosslink, whereas the other reveals a strikingly distorted DNA structure and a relaxed crosslink. Namely, the distorted duplex contains a significant bend and a nearly broken **G**:C base pair, with the M.*Hae*III substrate cytosine being partially extruded from the helix. Comparison of these two structures leads to a description of how and why torsional pressure applied from the minor groove can lead to rupture of base pairing in the major groove.

Materials and Methods

Sample Preparation and Crystallization. The palindromic oligonucleotide d(CCAGGCCTGG) was synthesized by using solid phase synthesis, crosslinked by using -CH_2-CH_2-S tethers attached to the N2 atom of the G5 bases on both strands, and purified by reversed phase HPLC and PAGE, as described previously (14, 15). The crosslinked duplex was dissolved in

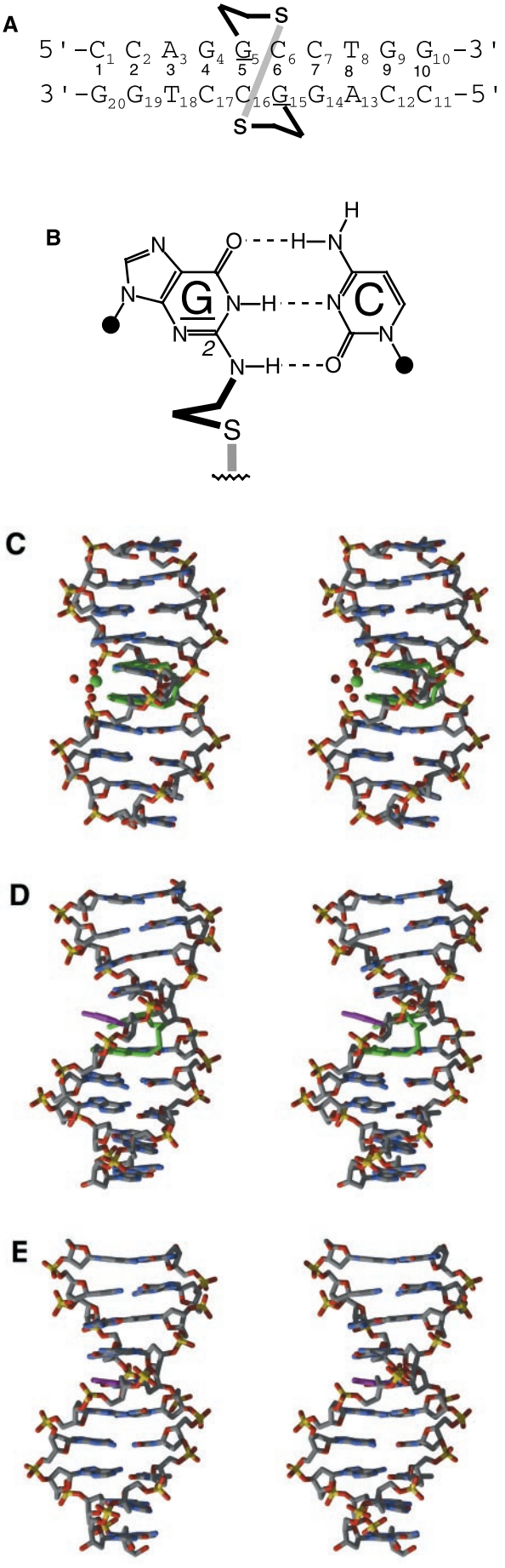
Abbreviation: rmsd, root-mean-square deviation.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 1CW9).

§To whom reprint requests should be addressed. E-mail: leemor@cshl.org.

‡Present address: Sunesis Pharmaceuticals, 3696 Haven Avenue, Suite C, Redwood City, CA 94063.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



water to a concentration of 1 mM and crystallized by using hanging-drop vapor diffusion. A matrix screen was used to find initial conditions that were subsequently optimized. Equal volumes of DNA solution and a solution containing 5% 2-methyl-2,4-pentanediol (MPD), 200 mM $CaCl_2$, 80 mM NaCl, 12 mM spermine tetrachloride, and 50 mM cacodylate, pH 7.0, were mixed on glass coverslips and sealed onto wells containing 35% MPD. Crystals appeared within 1 week at 4°C, to a maximum size of approximately 0.1 × 0.1 × 0.2 mm.

Data Collection, Structure Determination, and Refinement. Diffraction data were collected to 1.55 Å resolution on beamline X26C at the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY), by using a 300-mm MAR Research (Hamburg, Germany) image plate detector. Data were collected at 100 K by using cryocrystallographic techniques from a 0.05 × 0.05 × 0.1-mm crystal in an Oxford Cryosystems (Oxford, U.K.) cryostream. Diffraction data were integrated and scaled with the HKL package (18). The structure was solved by molecular replacement with AMORE (19) with the unmodified d(CCAGGCCTGG)$_2$ structure [Nucleic Acid Database entry BDJ017 (20)] as a search model. Unit cell dimensions indicated two duplexes per asymmetric unit, duplex A and duplex B. A clear solution was found by using 8.0–4.0 Å data, which after rigid body refinement in AMORE gave a crystallographic *R*-factor of 0.46. Refinement was performed with X-PLOR (21) and CNS (22), using the maximum likelihood function (23). The free-*R* test (24) was applied throughout the refinement with 5% of the reflections. The refinement was further monitored by studying the density of the crosslinks which were not included in the initial model. No σ cutoff was applied. Low-resolution data were included using the standard CNS bulk solvent correction (22). After several rounds of model building with O (25) and refinement (13.0–1.85 Å, R = 0.30, R_{free}= 0.32), both crosslinks and a strong spherical coordinated density were observed in $F_{obs} - F_{calc}$ and $2F_{obs} - F_{calc}$ maps, contoured at 2.5σ and 1.0σ, respectively. Initially, this peak, located in the major groove, was modeled as a water molecule, but extra positive density was observed in the maps and the B-factor refined to 3 Å^2. After inclusion of some additional water molecules (4σ peaks in an $F_{obs} - F_{calc}$ map and at least one hydrogen bond to a DNA nitrogen or oxygen atom), it became apparent that the strong peak was hepta coordinated, as found for Ca^{2+} ions in some nucleic acid structures (26). After modeling the peak as a Ca^{2+} ion, the positive density disappeared (and no negative density appeared), density of the coordinating ligands improved, and the B-factor assumed a more reasonable value (17 Å^2). After another round of model building, refinement, and inclusion of additional water molecules, density of the crosslinks improved and was continuous from G5-N2 on one strand to G5-N2 of the other strand (see Fig. 3). At this stage, coordinates for the crosslinks were included (13.0 − 1.55 Å, R = 0.251, R_{free}= 0.253). Interpretation of the maps allowed placement of more water molecules. At several stages during the refinement, omit maps were calculated for

Fig. 1. (*A*) Sequence of the C_2-crosslinked DNA decamer used in this study. The underlined G nucleotides denote the crosslinked guanines. (*B*) Chemical structure of a base pair containing the tether on the amino group in the minor groove of guanine. Stereo images of duplex A (*C*) and duplex B (*D*). The crosslinked guanines and the crosslink bridges are colored green. In duplex A, the calcium ion (green) and its liganding water molecules (red) are shown. In duplex B, the partially flipped-out cytosine (C6) is shown in magenta. (*E*) A canonical B–DNA duplex shown in the same orientation as duplex B with the cytosine corresponding to the partially flipped-out one, C6 in duplex B, shown in magenta for reference. By comparison, it is clear how much C6 is swung out and into the major groove. Fig. 1 *C–E* as well as Figs. 3 and 4 were prepared with the programs BOBSCRIPT (38), MOLSCRIPT (39), and RASTER 3D (40, 41).

individual or pairs of bases. A standard CNS-simulated annealing protocol was used, heating to 3,000 K with subsequent cooling, to eliminate bias. All omitted bases could be observed in the calculated $F_{obs} - F_{calc}$ maps, including the partially flipped-out base, which is particularly well ordered. The validation program SFCHECK (27) was used as an additional assessment of the final model. Further details on data collection and refinement statistics are given in Table 1.

Dynamics Simulation and Analysis. Dynamic properties of native and crosslinked d(CCAGGCCTGG)$_2$ were studied by computer simulation with the GROMOS suite of programs (28). Four different simulations were performed. One was started from the native d(CCAGGCCTGG)$_2$ crystal structure (20) (NAT). The second was started from the same structure with the crosslink from duplex B modeled in (NATXLB). Another started from duplex B (C2XB), and finally a simulation was started from the duplex A structure described here (C2XA). In all simulations, the DNA duplex was placed in a rectangular box filled with water molecules (29) (2,440 waters for NAT, 2,447 for NATXLB, 2,200 for C2XB, and 2,250 for C2XA). The system was then energy minimized by using the conjugate gradients algorithm until no significant energy change was observed. A pairlist cutoff of 8 Å was applied, together with a 12-Å long-range electrostatic cutoff. Backbone phosphates had no net charge, but were highly polar (charge of 1.44 on phosphorus −0.36 on oxygen atoms). No explicit counterions were used. The energy minimized structures were subjected to a 1-ps startup molecular dynamics run, for which initial velocities were taken from a Maxwellian distribution at 300 K. A 500-ps production run was performed, with SHAKE to constrain bond lengths (30), allowing a time step of 2 fs. The same distance cutoffs were applied as during the energy minimization. Temperature (300 K) and pressure (1 atm) coupling were applied (31) by using coupling constants of 0.1 ps and 0.5 ps, respectively. Trajectory frames were stored at a rate of 10/ps, and the last 400 ps were taken to represent equilibrium as judged by root-mean-square deviation (rmsd) with respect to the starting structures. Large concerted fluctuations of atoms were analyzed by using the essential dynamics method (32).

Results

Two Different Duplexes. The crystal structure of crosslinked d(CCAG**G**CCTGG)$_2$ contains two independent duplex molecules, which differ in several key aspects. Both types of duplexes form pseudocontinuous rods throughout the crystal (not shown). Visual inspection of the duplexes (Fig. 1) reveals that they possess distinct overall structures, the most obvious being that duplex A is nearly straight, whereas duplex B is significantly bent. While superposition of all P atoms of the two duplexes gives a rmsd of 1.58 Å, overlay of only the lower and upper five base pairs yields rmsds of 0.75 Å and 0.78 Å, respectively, thus indicating that the two duplexes deviate most substantially from one another in the middle region. Despite the formal dyad symmetry of the crosslinked sequence, neither duplex A nor B possesses identical half-site structures indicative of C2 symmetry. Duplex A exhibits helical parameters (not shown) that are similar to the native d(CCAGGCCTGG)$_2$ structure (20) with a rmsd of 1.03 Å, except that it contains a partially hydrated calcium ion directly coordinated to the crosslinked guanines in the major groove. On the other hand, the helical parameters of duplex B deviate substantially from the native structure (rmsd 1.34 Å), especially over the four middle base pairs (bp 4–7) on which the crosslink is centered. Both base pairs five and six of duplex B show a large roll (≈10°) (Fig. 2), sharply bending the helix axis mostly toward the major groove. The bending in duplex B is accompanied by substantial widening of the minor groove (7.9–9.0 Å in B, compared with 5.9 Å for canonical B-DNA and 5.1–7.0 Å in A).

The Ca^{2+} Ion. The presence of a Ca^{2+} ion coordinated to the O6 atoms of the crosslinked guanines was apparent in the electron-density maps at 1.55 Å (Fig. 3). Both O6-Ca^{2+} distances are 2.5 Å, indicative of strong inner-sphere coordination; the remaining coordination sites on the metal are occupied by five water molecules, resulting in a distorted octahedral coordination geometry. The Ca^{2+}-bound guanines appear slightly rolled toward the ion, perhaps permitting them to attain a more favorable coordination. Hepta-coordinated Ca^{2+} ions with water molecules as ligands have been observed previously in DNA crystal structures (17, 33–35), but to our knowledge this is the first

Table 1. Data collection and refinement statistics

Data collection	
Temperature, K	100
Wavelength, Å	1.08
Space group	$P2_12_12_1$
Cell dimensions, Å	$a = 34.44$, $b = 43.18$ and $c = 72.41$
Resolution range, Å	13.0−1.55 (1.61−1.55)
Number of observed reflections	62,990
Number of unique reflections	15,570 (1,555)
Redundancy	4.0
$I/\sigma I$	9.0 (3.0)
Completeness, %	95.5 (98.0)
R_{merge}, %	7.8 (32)
Refinement statistics for the final model	
Number of atoms	820 DNA, 1 Ca^{2+}, 271 water molecules
R, R_{free} (0.0 σ cutoff), %	20.3, 23.2
R, R_{free} (2.0 σ cutoff), %	19.5, 22.6
rmsd bonds, Å	0.009
rmsd angles, °	1.53
Overall average B factors, Å^2	17.4, 19.1 (duplex A, B)
Phosphate average B factors, Å^2	22.1, 24.3
Sugar average B factors, Å^2	18.1, 20.4
Base average B factors, Å^2	14.7, 15.8

Values in parentheses are for the highest-resolution shell.

BIOCHEMISTRY

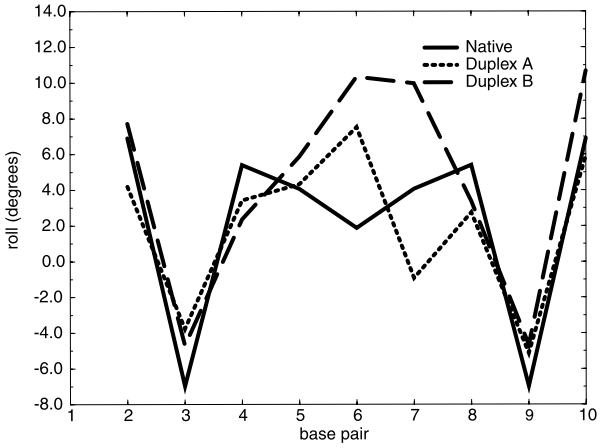


Fig. 2. Roll angles for the native structure (20), duplex A, and duplex B (this work). These parameters were calculated by using the program SCHNAAP (42).

instance of Ca^{2+} being simultaneously coordinated to adjacent base pairs of a B-DNA helix.

The Crosslinks. The crosslinks were included only in the later stages of structure refinement, once their electron density was unequivocal in difference maps. Even though both crosslinks have an S-S torsion angle close to the optimal value, $\phi_{S1S2} = 90^o$ (Fig. 3), the overall conformations of the tethers are markedly different (Fig. 3*C*). Although the DNA in molecule A is relaxed, its tether appears strained; specifically, one of the two -CH_2CH_2- units in the tether adopts a nearly eclipsed conformation, with three steric interactions, about 3 kcal/mol less stable than the staggered conformation of the other ethano bridge in the molecule A tether (36, 37). It is also very tightly packed into the minor groove. Importantly, this eclipsing interaction is relieved in molecule B, with only a single steric interaction around one of the CH_2-S bonds. In this case, the tether is loosely packed in the groove. It should be noted that the eclipsed bond in molecule A, as well as the relaxed crosslinked tether in molecule B, appears as continuous density in the unbiased difference electron density maps before their inclusion in the model being refined (Fig. 3 *A* and *B*). Numerous additional torsion angle adjustments accompany relief of the eclipsing interaction, the net result being that in going from molecule A to B, the tether is elongated and thrust forward in the direction of the bases by ≈2.5 Å (Fig. 3*C*), as if it were a tight spring that was released. By virtue of these changes, the entire disulfide crosslink in molecule B arrives at one of the most stable conformations capable of connecting the two **G**-N^2 atoms, loosely packed in the groove and having an extended (*trans*) ethano bridge connecting the disulfide to the guanine (**G**15) that is pushed into the major groove.

A Ruptured Base Pair. The consequence of this tether elongation pushing **G**15 of molecule B into the major groove is surprising. From the early stages of refinement, it was clear that C6 of the C6·**G**15 base pair of duplex B did not possess normal Watson–Crick hydrogen-bonding interactions and was not stacked conventionally with the surrounding bases. Indeed, the final model reveals that this base pair is partially ruptured, with an intact O2-N2 hydrogen bond (2.8 Å) on the minor groove side, but broken N3-N1 (4.8 Å) and major groove N4-O6 hydrogen bonds (7.0Å) (Fig. 3*B*). This mode of disruption is achieved by antiphase rotation of the bases in the plane of the ring about the helix axis by ≈60°. The space opened up by separation of the bases in the major groove is occupied by a water molecule, which appears to be hydrogen-bonded to the O6 and N1-H of **G**15 and to the N3 of C6. Whereas **G**15 is rolled over and almost completely unstacked from C16, it forms a nearly normal stacking interaction with G14, and the helix then continues normally on this side; the abrupt dislocation of the helix between **G**15 and C16 is responsible for the observed bending of the helix. Even more striking is the disposition of C6, which is partially flipped out of the helix (Fig. 3*B*). This cytosine is in fact stabilized in this extrahelical conformation by a crystal hydrogen-bonding contact between its exocyclic N4-H and the nonbridging phosphate oxygen of G20 borne on a symmetry-related duplex molecule A. The space that would ordinarily have been occupied by the O2 atom of a regularly stacked cytosine and N2 of guanine has been partially filled by the atoms of the crosslink. As mentioned earlier, the oligonucleotide used in this study is the target sequence for *Hae*III methylase. Remarkably, this partially flipped cytosine is indeed the same cytosine that is targeted by the enzyme for extrusion, insertion into the active site, and transfer of a methyl group.

Dynamics. To examine the mobility of the partially extrahelical cytosine in solution, we compared the molecular dynamics trajectories of the central two base pairs in unmodified and crosslinked DNA contained in a box of water molecules. From the 1-ns trajectories, only the coordinates for base pairs 5 (**G**5-C16) and 6 (C6-**G**15) were extracted. These were then analyzed by using essential dynamics (32), by building a covariance matrix of the native and a crosslinked DNA trajectory together (8,000 frames). Three such matrices were built, one for the unmodified native duplex (NAT) with duplex B of our structure (C2XB), one for NAT with a native duplex that has the crosslink from duplex B of our structure modeled in (NATXLB) and another for NAT with duplex A of our structure (C2XA). The eigenvectors of these "combined" covariance matrices describe motions that are present in both simulations. We projected the separate 1-ns trajectories from the comparisons onto the combined eigenvectors to study differences in dynamic behavior. In the NAT + C2XB analysis, two eigenvectors, the second and third, show significant differences. The eigenvector with the third largest eigenvalue (i.e., the third largest concerted motion in this two base pair system) shows about three times more fluctuation in native DNA. Visual inspection of the fluctuation described by this eigenvector shows a twisting motion of the base pairs in opposite directions around the helical axis. This kind of motion, which would require a separation of the N2 atoms on the crosslinked bases, is significantly inhibited by introduction of the crosslink. The motion described by the second eigenvector also describes a difference in dynamic behavior between the native and crosslinked forms (Fig. 4). Here, the flipped-out cytosine is moving back and forth in the base pair plane. The amplitude of this motion is twice as large in the crosslinked DNA, compared with native DNA. From this analysis, it appears that introduction of the crosslink may lead to increased mobility of the target cytosine, which is base paired to the crosslinked guanine. Analysis of NAT + NATXLB, which was performed as a control to exclude structural bias in comparing the dynamics, showed similar results: an increased flipping motion of the cytosine, with a more restrained backbone. Another control, the native with duplex A (NAT + C2XA), showed some increased motion of the cytosine in duplex A that was much reduced compared with duplex B. Indeed, the conformation of the crosslink in this case remains close to the original tight eclipsed conformation for this 1-ns simulation.

Discussion

Here we have captured, in the same crystal, two very different conformational states of a single DNA duplex presented with the challenge of accommodating a disulfide crosslink. In the state represented by duplex A, the structure of the DNA helix is so

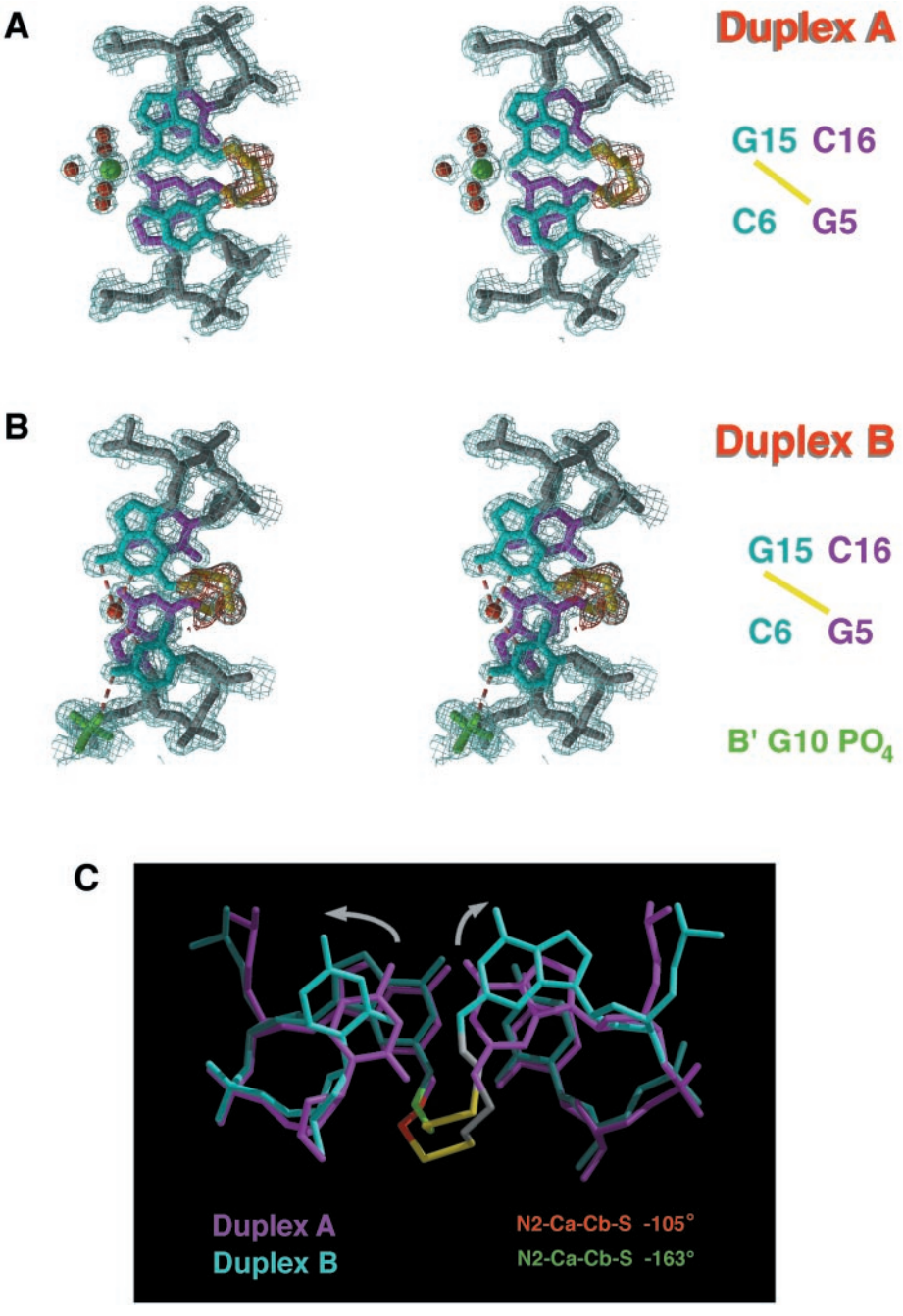


Fig. 3. Stereo views of the structure of the middle two base pairs, the crosslinks (yellow), and the calcium (green) ion for duplexes A (*A*) and B (*B*). Also shown is the phosphate (green) on a symmetry related duplex B, which interacts with the partially flipped-out cytosine of duplex B. Final $1.0\sigma\, 2F_{obs} - F_{calc}$ maps are shown in light blue. The $2.5\sigma\, F_{obs} - F_{calc}$ map just before including the crosslinks in the model is shown in red. Water molecules discussed in the text are shown as red spheres. Hydrogen bonds (under 3.5 Å) are drawn for the water in duplex B. (*C*) A superposition of the two crosslinked base pairs from duplex A in magenta and duplex B in cyan with the crosslinks in gray and yellow for the sulfurs. The nearly eclipsed bond in the A crosslink is shown in red, and the same bond in its staggered conformation in the B crosslink is shown in green. The corresponding torsion angles are noted as well. The arrows depict the swiveling of the bases in duplex B from a standard conformation in A.

unperturbed as to be nearly indistinguishable from that of the corresponding unmodified duplex. This native-like structure of duplex A stands in marked contrast to the helical structure of duplex B, which bears striking aberrations, including a significant bend toward the major groove, a partially ruptured base pair, and a cytosine residue that is well along the way toward being extruded from the helix. Despite these differences in helical structure, the fact that duplexes A and B are observed together in the same crystal suggests that they are roughly equivalent in terms of overall energy. These considerations point to the existence of a seemingly cryptic source of strain energy in duplex A, the obvious candidate for which is the disulfide crosslink itself. Indeed, inspection of the disulfide crosslink in duplex A reveals that it is tightly packed in the minor groove and that it contains an eclipsed C-C bond. An eclipsed bond is a highly disfavored interaction that releases more than 3 kcal/mol upon bond rotation to yield a staggered conformation. In going from molecule A to B, the most significant change in tether confor-

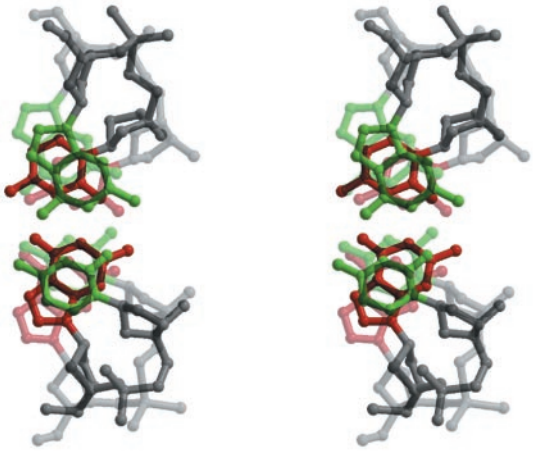


Fig. 4. Stereo images of snapshots of the motion along the second eigenvector. Two structures are shown: two extremes of the motion [projection −0.75 nm (transparent) and projection 0.75 nm (opaque)]. The bases in the green base pair show an anticorrelated motion, breaking up the hydrogen bonds and partially swinging out the cytosine.

mation is rotation of the eclipsed bond by ≈60° to yield the staggered *trans* conformation. As the C-C bond rotates to the *trans* conformation, the tether becomes substantially elongated on one side (Fig. 3*C*) and more loosely packed in the groove. The tether in molecule A can be envisioned as a tight spring, held tight perhaps with the help of the calcium ion on the opposite groove. It is then released into a longer more relaxed conformation in molecule B. This thrusts the N2-atom of the attached guanine base forward into the major groove. However, because one end of this guanine base remains attached to the DNA backbone, the base is constrained to swivel rather than slide forward. Although a sliding motion would permit retention of Watson–Crick base pairing, swiveling causes the base pair to separate in the major groove. The minor groove hydrogen-bonding partners remain near enough to retain hydrogen bonding, perhaps only transiently should the cytosine completely flip out. This conformation is subsequently trapped by a stabilizing hydrogen-bonding interaction with a neighboring molecule. Thus, the net result of this process is to generate molecule B directly. We therefore propose that duplexes A and B are the products of internal transfer of strain energy in the crosslinked decamer, with the strain principally residing in the tether (A) or DNA helix (B).

It is tempting to speculate that certain proteins might use a mechanism like the one elucidated here to induce the breakage of base pairs in DNA. Our results suggest that a "push" of only ≈3 kcal/mol from the minor groove suffices to break a G·C base pair and exert the cytosine residue from the DNA helix. This relatively modest amount of energy is readily attainable through noncovalent intermolecular interactions between the DNA and the enzyme paralleling its stabilization by intermolecular interactions with a neighboring DNA molecule in the crystal we describe here. As with the disulfide crosslink, the strain energy generated upon binding of a protein to DNA can be transduced to the helix through a remarkably simple process, resulting in relaxation of the protein–DNA interface and concomitant rupture of a base pair.

We thank Malcolm Capel for help with data collection, Dorothy Wang with early crystallization experiments, and the members of the L.J. group for fruitful discussions. This research was supported by grants of the Arnold and Mable Beckman Foundation and the Robertson Research fund (both to L.J.) and by a long-term European Molecular Biology Organization fellowship (D.v.A.).

1. Joshua-Tor, L., Frolow, F., Appella, E., Hope, H., Rabinovich, D. & Sussman, J. L. (1992) *J. Mol. Biol.* **225,** 397–431.
2. Morden, K. M., Chu, Y. G., Martin, F. H. & Tinoco, I., Jr. (1983) *Biochemistry* **22,** 5557–5563.
3. Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F. & Patel, D. J. (1989) *Biochemistry* **28,** 294–303.
4. Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F. & Patel, D. J. (1990) *J. Biol. Chem.* **265,** 636–647.
5. Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. D. (1994) *Cell* **76,** 357–369.
6. Reinisch, K. M., Chen, L., Verdine, G. L. & Lipscomb, W. N. (1995) *Cell* **82,** 143–153.
7. Lau, A. Y., Schärer, O. D., Samson, L., Verdine, G. L. & Ellenberger, T. (1998) *Cell* **95,** 249–258.
8. Vassylyev, D. G., Kashiwagi, T., Mikami, Y., Ariyoshi, M., Iwai, S., Ohtsuka, E. & Morikawa, K. (1995) *Cell* **83,** 773–782.
9. Savva, R., Mcauley-Hecht, K., Brown, T. & Pearl, L. (1995) *Nature (London)* **373,** 487–493.
10. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E. & Tainer, J. A. (1996) *Nature (London)* **384,** 87–92.
11. Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J. & Pearl, L. H. (1998) *Cell* **92,** 117–129.
12. Roberts, R. J. & Cheng, X. (1998) *Annu. Rev. Biochem.* **67,** 181–198.
13. Wolfe, S. A., Ferentz, A. E., Grantcharova, V., Churchill, M. E. A. & Verdine, G. L. (1995) *Chem. Biol.* **2,** 213–221.
14. Erlanson, D. A. & Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 12583–12584.
15. Erlanson, D. A., Wolfe, S. A., Chen, L. & Verdine, G. L. (1997) *Tetrahedron* **33,** 12041–12056.
16. Wolfe, S. A. & Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 12585–12586.
17. Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J. & Dickerson, R. E. (1993) *Biochemistry* **32,** 1373–1389.
18. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
19. Navaza, J. & Saludjian, P. (1997) *Methods Enzymol.* **276,** 581–594.
20. Heinemann, U. & Alings, C. (1989) *J. Mol. Biol.* **210,** 369–381.
21. Brünger, A. T. (1988) *J. Mol. Biol.* **203,** 803–816.
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Gross-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr.* **D 54,** 905–921.
23. Pannu, N. S. & Read, R. J. (1996) *Acta Crystallogr.* **A 52,** 659–668.
24. Brünger, A. T. (1992) *Nature (London)* **355,** 472–474.
25. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr.* **A 47,** 110–119.
26. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R. & Schneider, B. (1992) *Biophys. J.* **63,** 751–759.
27. Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999) *Acta Crystallogr.* **D 55,** 191–205.
28. Gunsteren, W. F. v. & Berendsen, H. J. C. (1987) *Biomolecular Software* (University of Groningen, The Netherlands).
29. Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F. v. & Hermans, J. (1981) in *Intermolecular Forces*, ed. Pullmann, B. (Reidel, Dordrecht).
30. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977) *J. Comp. Phys.* **23,** 327–341.
31. Berendsen, H. J. C., Postma, J. P. M., DiNola, A. & Haak, J. R. (1984) *J. Chem. Phys.* **81,** 3684–3690.
32. Amadei, A., Linssen, A. B. M. & Berendsen, H. J. C. (1993) *Proteins* **17,** 412–425.
33. Yuan, H., Quintana, J. & Dickerson, R. E. (1992) *Biochemistry* **31,** 8009–8021.
34. Grzeskowiak, K., Goodsell, D. S., Kaczor-Grzeskowiak, M., Cascio, D. & Dickerson, R. E. (1993) *Biochemistry* **32,** 8923–8931.
35. Goodsell, D. S., Grzeskowiak, K. & Dickerson, R. E. (1995) *Biochemistry* **34,** 1022–1029.
36. Lide, D. R. (1958) *J. Chem. Phys.* **29,** 1426–1427.
37. Weiss, S. & Leroi, G. E. (1968) *J. Chem. Phys.* **48,** 962–967.
38. Esnouf, R. M. (1997) *J. Mol. Graphics* **15,** 132–134.
39. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
40. Bacon, D. J. & Anderson, W. F. (1988) *J. Mol. Graphics* **6,** 219–220.
41. Merritt, E. A. & Murphy, M. E. P. (1994) *Acta Crystallogr.* **D 50,** 869–873.
42. Lu, X. J., Hassan, M. A. E. & Hunter, C. A. (1997) *J. Mol. Biol.* **273,** 668–680.