Structure determination of DNA methylation lesions
\(N^1\)-meA and \(N^3\)-meC in duplex DNA using a
cross-linked protein–DNA system

Lianghua Lu, Chengqi Yi, Xing Jian, Guanqun Zheng and Chuan He*

Department of Chemistry, The University of Chicago, 929 E. 57th Street, Chicago, IL 60637, USA

Received December 18, 2009; Revised February 11, 2010; Accepted February 12, 2010

ABSTRACT
\(N^1\)-meA and \(N^3\)-meC are cytotoxic DNA base
methylation lesions that can accumulate in the
genomes of various organisms in the presence of
\(S_2\) type methylating agents. We report here the
structural characterization of these base lesions in
duplex DNA using a cross-linked protein–DNA crys-
tallization system. The crystal structure of \(N^1\)-meA:T
pair shows an unambiguous Hoogsteen base pair
with a syn conformation adopted by \(N^1\)-meA,
which exhibits significant changes in the opening,
roll and twist angles as compared to the normal
A:T base pair. Unlike \(N^1\)-meA, \(N^3\)-meC does not
establish any interaction with the opposite G,
but remains partially intrahelical. Also, structurally
characterized is the \(N^6\)-meA base modification that
forms a normal base pair with the opposite T in
duplex DNA. Structural characterization of these
base methylation modifications provides molecular
level information on how they affect the overall
structure of duplex DNA. In addition, the base
pairs containing \(N^1\)-meA or \(N^3\)-meC do not share
any specific characteristic properties except that
both lesions create thermodynamically unstable
regions in a duplex DNA, a property that may be
explored by the repair proteins to locate these
lesions.

INTRODUCTION
Genomic DNA is constantly subjected to modifications
caused by exogenous environmental chemicals and
cellular metabolites. Among various DNA damages,
nucleobase damage is a common type that can be
induced by oxidation, hydrolysis and alkylation (1,2). If
left unrepaired, these lesions can cause deleterious effects
on nearly all aspects of cellular functions (3,4). Therefore,
guarding genetic integrity by repairing DNA damage is
one of the most fundamental processes of life (5,6). Most DNA base damage is repaired through four
biochemically and mechanistically distinct pathways:
direct reversal repair (DRR), base excision repair (BER),
nucleotide excision repair (NER) and mismatch repair
(MMR) (7).

Structural characterization of base lesions in
double-stranded DNA (dsDNA) is important to under-
stand the origin of their mutagenic or cytotoxic effects.
The structural information may also help to reveal how
the lesions are detected and repaired by DNA repair
proteins. Despite considerable advances in DNA structure
characterization (8), high-quality crystals of naked
dsDNA containing base lesions are hard to obtain. Even
with the Dickerson–Drew dodecamer sequence (9–12),
base lesions can introduce unstable base pairs, which
prevent the formation of well-ordered 3D crystal lattice.
In order to aid DNA crystallization, host–guest systems
employing various DNA-binding proteins have been used
to obtain high-quality crystals (13–17). Although this
method has helped to solve the structures of some impor-
tant lesions (16,17), many more remain to be elucidated.
The host–guest approach has also been shown to be effec-
tive for RNA crystallization (18–22).

Alkylated nucleobases are produced mostly by
endogenous and environmental alkylation agents (23).
\(N^1\)-methyladenine (\(N^1\)-meA) and \(N^3\)-methylcytosine
(\(N^3\)-meC) are two forms of base methylations generated
usually in single-stranded DNA (ssDNA). These lesions
are unable to form normal Watson–Crick base pairs,
and consequently block DNA replication, resulting in
cytotoxicity to cells. Escherichia coli \(alkB\) is an inducible
gene of the adaptive response to alkylating agents. The
\(alkB\) gene encodes a DNA-repair protein that counteracts
toxic alkylation damages, such as \(N^1\)-meA, \(N^3\)-meC and
1,\(N^6\)-ethenoadenine (24–28). Human homologues of \(alkB\)
can process similar base modifications and exhibits
profound functional roles in human cells (29–31).

Solving the structures of these base modifications will be
important to visualize how these lesions affect the local dsDNA structure, thereby shedding light on the origin of the cytotoxic effects of the damage. The structure of 1, N⁶-ethenoadenine:T pair was recently characterized using the DNA glycosylase AlkA bound to the ends of dsDNA (17). The crystal structures of the other two lesions, N¹-meA and N³-meC in dsDNA, remain unknown. We have solved the structure of hABH2, the human protein that repairs N¹-meA and N³-meC in ssDNA and dsDNA (32,33), cross-linked to dsDNA (34). As most of the DNA bases have no direct contact with the protein residues in the ABH2–dsDNA structure (PDB 3BTX, Supplementary Figure 1), this complex seems to be a suitable host–guest-like system for characterizing lesion-containing dsDNA structures. In our design, the protein–dsDNA complex is stabilized through covalent cross-linking in the protein active site, whereas the lesioned base is installed to the locations without any direct contact to the protein (Supplementary Figure S2). A noticeable advantage of this approach versus other host–guest-like systems is its robustness since the complex is stitched together with a covalent bond. The complex cannot be crystallized in the absence of this cross-linking.

This host–guest-like system is validated through the structural characterization of the methylation base modifications N¹-meA, N³-meC and N⁶-meA in dsDNA cross-linked to ABH2. In these structures, N¹-meA forms an unambiguous anti-Hoogsteen base pair to the opposite T; N³-meC does not form any hydrogen-bonding interaction with the opposite G, but stays partially intrahelical inside the duplex DNA.

MATERIALS AND METHODS

Oligonucleotide synthesis and purification

Oligonucleotides containing disulfide-tethered cytosine and specific base lesions were synthesized using the phosphoramidite derivative of O⁴-triazoyl-dU (Glen Research), and functionalized with cystamine. The DNA oligomers were synthesized on a Biosystems Expedite Nucleic Acid Synthesis System using standard reagents. Following synthesis, the resin was dried under vacuum and then incubated with 200 µl of solution containing 50% cystamine in water at room temperature for 16 h. The aqueous solution was collected and the resin was washed twice with 200 µl of water. The washes were combined and neutralized to pH 6.0–6.5 with glacial acetic acid on ice and desalted by a NAP-10 column (GE healthcare). The aqueous solution was then lyophilized and finally purified by reverse-phase HPLC. The dsDNA used in the crystallization of N¹-meA and N⁶-meA is DNA1: 5'-CTGCTATC*AT(mA)GCG-3' paired with 5'-TGGCTATAATACA-3'. The dsDNA used for N³-meC is DNA2: 5'-CTGT(mG)TC*ATTGC G-3' paired with 5'-TCGCAATAAGACA-3'. Masses of the synthesized oligonucleotides were verified by MALDI-TOF mass spectrometry.

Cloning, expression and purification of ABH2–ΔN55 E175C

The abh2ΔN55 gene was cloned between the NdeI and HindIII sites of a pET28a vector (Novagen). ABH2–ΔN55 E175C mutants were generated by using QuickChange Site-Directed Mutagenesis Kit (Stratagene) and this plasmid was transformed into E. coli BL21(DE3) cells (Stratagene) for protein overexpression. The protein was purified using HisTrap HP column (GE Healthcare) with an elution buffer of 50 mM sodium phosphate (pH 8.0), 300 mM NaCl and 400 mM imidazole. The N-terminal His-tag was removed by an overnight thrombin (MP Bio) digestion at 4°C and the tag-free protein was purified the next day with another round of HisTrap chromatography. About 4-6 mg of protein could be obtained from 11 of bacterial culture.

Preparing cross-linking, crystallization and data collection of the ABH2–DNA complexes

DNA duplexes were annealed by mixing 1 mM thiol-tether containing strand with the corresponding complimentary strand in buffer containing 10 mM Tris (pH 7.4) and 100 mM NaCl, incubating at 75°C for 10 min and cooling to 4°C by a step gradient of −1°C/min. The cross-linked complexes of protein ABH2–ΔN55 E175C with synthetic oligonucleotides were achieved by incubating dsDNA (1 mM, 50 µl) with protein (1.2 eq.) in 5 ml of buffer [100 mM NaCl, 10 mM Tris–HCl (pH 7.4)] at 4°C for 16 h. The covalently linked ABH2–DNA complexes were purified using MonoQ anion exchange chromatography (GE healthcare), then buffer exchanged to a solution containing 100 mM NaCl, 10 mM Tris–HCl (pH 8.0) (Supplementary Figure S3). Finally, the complexes were concentrated to 5 mg/ml for use in crystallization experiments.

Protein–dsDNA complex crystals were grown by hanging drop vapor diffusion crystallization at 4°C in drops containing 1 µl of complex solution and 1 µl of reservoir solution of 100 mM NaCl, 50 mM MgCl₂, 100 mM cacodylate (pH 6.5) and 12% 5 K PEG. Hexagonal rod-shaped crystals grew in 1–2 weeks at 4°C. Subsequently, crystals were transferred to a cryoprotectant solution composed of 80% reservoir solution and 20% glycerol and frozen in liquid nitrogen before data collection. An X-ray data set (diffracted to 2.0 Å) for the N¹-meA-containing complex (ABH2–N¹-meA) crystals and an X-ray data set (diffracted to 1.8 Å) for the N⁶-meA-containing complex (ABH2–N⁶-meA) were collected at beamline 23ID-B (General Medicine and Cancer Institutes Collaborative Access Team [GM/CA-CAT]) of the Advanced Photon Source at Argonne National Laboratory. The N³-meC-containing complex (ABH2–N³-meC) gave the best crystals that diffracted to 2.0 Å (Table 1). Data were integrated and processed with the HKL2000 package.

Structure determination and refinement

The ABH2–dsDNA complex structures were phased by molecular replacement [using Phaser (35)], using the
previously published ABH2–dsDNA structure as a search model [PDB 3BTX (34)]. The model was built by using COOT and refinement was carried out with the program REFMAC5 from the CCP4 suite. Atomic coordinates for the three structures reported have been deposited in Protein Data Bank under accession number 3H8O (ABH2–N1-meA), 3H8R (ABH2–N6-meA) and 3H8X (ABH2–N3-meC).

RESULTS

Overall structures

In an attempt to crystallize duplex DNAs containing the methylated base N1-meA and N3-meC, we employed the Dickerson–Drew dodecamer sequence (9–12), which failed to yield crystals. Then, we decided to use the cross-linked ABH2-dsDNA complex (34) as a host–guest-like system for crystallization since the covalent cross-linking stabilizes the protein–dsDNA complexes. The lesion-containing dsDNA is cross-linked to the ABH2 protein with an engineered Cys at position 175 (ABH2–C1175C) (Supplementary Figure S2). Out of multiple DNA sequences screened, a 13-mer dsDNA1 (Figure 1B), with a central C*:A base pair (C* represents a disulfide-tethered cytosine), yielded the best quality crystals of the ABH2–dsDNA complexes containing N1-meA (Figure 1A) or N6-meA (Figure 1E). Another sequence, dsDNA2 (Figure 1D), with N3-meC at a different location, gave the best crystals for ABH2–dsDNA complex with N3-meC (Figure 1C). All of these complexes bearing modified nucleotides are crystallized in the hexagonal space group P622, and the crystals diffract to 1.8 Å for the N6-meA-containing complex (ABH2–N6-meA), and 2.0 Å for both N1-meA- and N3-meC-containing complexes (ABH2–N1-meA and ABH2–N3-meC, Table 1).

The structures of ABH2–N1-meA and ABH2–N6-meA overlap well with the original structure 3BTX (Supplementary Figure S5), with a root mean square deviation (RMSD) of 0.49 Å and 0.21 Å, respectively. Except for the site of the flipped-out C*:7, the DNA adopts a right-handed conformation and its interactions with the protein are mainly through the phosphodeoxyribose backbones, which are mostly located in the middle of the lesion-containing strand and 5′-end of the complementary strand. Interestingly, the structure of ABH2–N1-meA complex shows reduced contacts between protein residues and the backbone of dsDNA at the 5′-end of the complementary strand. For instance, the backbone region around T5 (which pairs with N1-meA) and its flanking nucleotides, A6 and C4′ (Figure 2), do not form hydrogen bonds with the protein residues Arg198, Gly204, Lys205 and Arg215; such hydrogen-bonding

### Table 1. Data collection and refinement statistics

| Crystal | ABH2–N1-meA:T | ABH2–N6-meA:T | ABH2–N3-meC:G |
|---------|---------------|---------------|---------------|
| Data quality | | | |
| Resolution (Å) | 20–2.0 (2.05–2.0) | 20–1.8 (1.77–1.8) | 20–2.0 (1.95–2.0) |
| Unique reflections | 29687 | 38371 | 3492 |
| Completeness (%) | 99.9 (99.9) | 99.9 (100) | 98.3 (78.6) |
| Redundancy | 20.0 | 20.3 | 21.1 |
| Rwork/σo (%) | 24.3 (2.6) | 32.6 (1.6) | 36.2 (2.2) |
| Crystal parameters | | | |
| Space group | P622 | P622 | P622 |
| Cell constants (Å) | | | |
| a = 79.11 | a = 79.11 | a = 79.11 |
| b = 79.11 | b = 79.11 | b = 79.11 |
| c = 242.27 | c = 228.71 | c = 242.30 |
| α = 90 | α = 90 | α = 90 |
| β = 90 | β = 90 | β = 90 |
| γ = 120 | γ = 120 | γ = 120 |
| Refinement | | | |
| Resolution (Å) | 20–2.0 | 20–1.77 | 20–1.95 |
| Rwork (%) | 20.2 | 21.0 | 20.5 |
| Rfree (%) | 23.7 | 23.6 | 23.6 |
| Model quality | | | |
| R.m.s. deviation bond (Å) | 0.015 | 0.010 | 0.0107 |
| R.m.s. deviation angle (°) | 1.58 | 1.53 | 1.57 |
| Average B factor | 27.7 | 29.2 | 36.5 |
| Average B factor of Protein | 27.1 | 27.9 | 34.6 |
| Average B factor of DNA | 23.1 | 29.0 | 39.6 |
| Average B factor of water | 39.4 | 42.5 | 45.3 |
| Model content | | | |
| Protein residues | 56–258 | 56–258 | 56–258 |
| Nucleotides | 26 | 26 | 26 |
| Water atoms | 197 | 222 | 179 |
| PDB accession code | 3H8O | 3H8R | 3H8X |

*Values in parentheses refer to the highest resolution bins.

Rwork = Σ |Fo–Fc|/Σ Fo, where Fo and Fc are observed and calculated structure factor amplitudes, respectively.

Rfree was calculated based on a percentage of data (5%) randomly selected and omitted through the structure refinement procedure.
contacts exist in the original structure of 3BTX. Besides Phe102, the finger residue that helps to flip C*7 out of the duplex DNA, no other protein residue affects base pairing in both structures of ABH2–N1-mea and ABH2–N6-mea as well as in 3BTX (34).

A superposition of the structure of ABH2–N3-mec to 3BTX shows an apparent DNA backbone shift toward the 3’-end of the complementary strand (Figure 2E), which is most likely caused by different packing of the DNA ends. The original structure (3BTX) has the terminal bases
twisted out of the duplex DNA to interact with the neighboring oligonucleotide; the terminal bases in ABH2–N³-meC remain intrahelical and form coaxial stacking with the neighboring duplex. As a result, the DNA bending induced by DNA end packing is slightly released when compared to the original structure (3BTX). Similar to ABH2–N¹-meA, a loss of contact between ABH2 and the backbone region around the nucleotides C⁴₀–A⁶₀ was observed in the structure of ABH2–N³-meC.

Structures of ABH2–N¹-meA and ABH2–N⁶-meA

Besides minor shift and rotation, the packing of the dsDNA helix in ABH2–N¹-meA is similar to the unlesioned DNA in 3BTX (Figure 3). The N¹-meA-containing dsDNA shows two slightly different backbone conformations, which do not seem to cause noticeable changes in the base-pairing mode of the DNA structure (Supplementary Figure S6). As N¹-meA can be converted to N⁶-meA via Dimroth rearrangement (36), we also crystallized ABH2–N⁶-meA for comparison. The DNA structure with N⁶-meA and its alignment to the unlesioned DNA in 3BTX are shown in Supplementary Figure S7. Besides an ~15° bending of the DNA caused by the end packing of DNA, the overall DNA conformation in these two structure remains basically B-form with an average rise per base pair along axis of around 3.2 Å, a slide of ~0.5 Å and a roll of 3.6°. Because N⁶-meA forms Watson–Crick base pair with the complementary T (Figure 4E), this N⁶-meA-containing dsDNA basically resembles an unmodified B-form-like duplex structure.
To perform a detailed analysis of the distortion of the duplex structure caused by the N1-meA lesion, the structure of the N1-meA-containing oligonucleotide was superimposed onto that of the N6-meA-containing DNA (Supplementary Figure S8). The base pair parameters of these two DNA structures were calculated using the program 3DNA v1.5 (37) and the changes for propeller twist, opening, buckle, roll and twist are shown in Figure 5. Major changes occur around the lesion site, where the positions of N1-meA10, complementary T5',...
and the flanking base pair G11:C4' shift relative to their positions in the structure of the N6-meA-containing duplex. These differences are results from the N1-methylation of A10, which blocks the normal Watson–Crick base pairing (Figures 3 and 4).

A Hoogsteen base pair between N1-meA10 and the opposite base T5' is clearly observed with two hydrogen bonds formed between these two bases (N2...O4, 3.1 Å; N7...N3, 3.2 Å; Figure 4B). Although the Hoogsteen base-pair mode has been reported by the NMR study of the N1-meA:T pair (38), the distances we observed here are much shorter than what have been assigned previously (around 4.3 Å). To accommodate both the methylated adenine and the pyrimidine ring of T within the duplex, the purine ring of N1-meA adopts a syn conformation with respect to the N9 atom around the N-glycosidic bond. In addition, a distinct dihedral angle of about 30° is found between the plane of N1-meA10 and the opposite T5' (Figures 4C and 6), which is significantly larger than that in the ABH2–N6-meA structure (Figure 4D).

Interestingly, despite dramatic changes occurring on the N1-meA10:T5' base pair, the neighboring base pair T9:A6' (5' to the lesion) is almost unaffected, but the base pair G11:C4' (3' to N1-meA), shifts slightly towards the 3' direction of the lesion-containing strand (Figure 6A). Compared to the N6-meA structure, the rise of the N1-meA:T pair to its flanking base pairs was decreased, which was caused by the inclination of the N1-meA base. In spite of this, the overall distance from T9:A6' base pair to G11:C4' is similar to that in a typical B-form DNA, which may indicate that the influence of the lesion is mainly limited to the local structure, but not the overall duplex.

Since the protein has no direct contact to the base pairing region around N1-meA10 or N6-meA10, an alignment of the two duplex DNAs in this region (Figure 6) allows a clear comparison of the properties of N1-meA with those of N6-meA, and reveals impacts on the local DNA structure by these modifications. It can be seen that N1-meA adopts a syn conformation to form the Hoogsteen base pair instead of the anti conformation adopted by N6-meA in a Watson–Crick base pair. A comparison of N1-meA:T pair with the N6-meA:T (Table 2) reveals that the shear of N1-meA:T pair is bigger (0.7 Å compared to 0.05 Å), and so is the stretch value (3.8 Å in N1-meA:T versus −0.2 Å in N6-meA:T). The stagger of N1-meA:T amounts to −1.1 Å (0.01 Å in N6-meA:T), which is mainly caused by a shift of the N1-methylated adenine into the major groove by −3.0 Å and a shift of the opposite T by 1.7 Å. The propeller twist of the methylated base pair is varied from 9.2° in N6-meA:T pair to 7.7° in N1-meA:T, while the two flanking base pairs are almost unaffected. The dramatically increased base opening of N1-meA:T pair is observed (73.6° compared to 0.3° in N6-meA:T, Figure 5B), as well as the greater buckle angle (29.1° versus 1.5°). However, the buckle value of G11:C4' in the N1-meA structure, 3' to the lesion pair, is decreased to −3.6°, which is 6.3° for the corresponding base pair in the N6-meA structure. Apparent changes are also observed in the N1-meA...
structure compared to the N6-meA and the unlesioned ones when calculating the base pair step parameters as shown in Table 2. The tilt and twist values of the N1-meA:T pair to both flanking base pairs show significant changes. With respect to the roll angles, only the value of the N1-meA:T pair to 30 adjacent G11:C40 pair is altered dramatically to /C0101.2/C14 (7.0 /C14 in N6-meA structure), but that of the 50-neighbored T9:A9 pair is only slightly affected. Furthermore, the C1–C1′ distance between N1-meA and the opposite T decrease to 9.0 A˚ (10.5 A˚ in N6-meA structure), and the distances of RN9–YN1 and RC8–YC6 are also shortened by 1.7 A˚ and 3.1 A˚ , respectively, compared to those in the N6-meA:T pair. These changes might be induced by the rotation of N1-meA base around the glycosidic bond, which ‘drags’ the complementary thymine further to the center of the helix (Supplementary Figure 9). A slight DNA distortion of the backbone is indicated by an elongation of the P_T9...P_A10 from 7.2 A˚ in the N6-meA-containing structure to 7.4 A˚ in the case of N1-meA, which probably explains the lack of interaction between the local backbone and the protein.

A comparison of the structure of N1-meA:T to that of the Hoogsteen base-paired A:T reveals that the same syn conformation of N1-meA or adenine is observed in both structures. A similar shortening of the C1–C1′ distance is also observed (9.0 A˚ for N1-meA:T pair, 8.5 A˚ for A:T Hoogsteen base pair, as opposed to 10.5 A˚ for A:T Watson–Crick base pair) (39,40). However, the distinct dihedral angle of /C2430/C14 between the plane of N1-meA10 and the opposite T50 as well as the major changes in twist angles in this region, are only found in the N1-meA:T structure, which could be explained by the steric effect of the N1-methylation.

Structure of dsDNA-containing N3-meC

Due to the lack of hydrogen bonding between N3-meC and the opposite G, the crystal structure of ABH2–N3-meC shows a distorted backbone compared to the complex without the base lesion. The DNA ends are packed differently in two structures (Figure 7) that helps to release the bending of the dsDNA caused by DNA end packing. The absence of any interaction between N3-meC to the opposite G leads to a shift of N3-meC towards the

Table 2. Comparison of base pair parameters of N1-meA:T, N6-meA:T and T:A

| Base Pairs          | Local base-pair parameters | Local base-pair step parameters |
|---------------------|----------------------------|---------------------------------|
|                     | Shear | Stretch | Stagger | Buckle | Propeller | Opening |
| N1-meA:T10          | 0.7   | −3.8    | −1.1    | 29.1   | 7.7       | 73.6    |
| N6-meA:T10          | −0.05 | −0.2    | 0.01    | 1.5    | −9.2      | 0.3     |
| T10:A10             | −0.05 | −0.2    | −0.2    | 1.4    | −21.0     | −0.5    |
| N1-meA:G11C         |       |         |         |        |           |         |
| N6-meA:G11C         |       |         |         |        |           |         |
| 3BTX                |       |         |         |        |           |         |

For shear, stretch and stagger, values are distances (A˚). Parameters for buckle, propeller and opening are angles (°).

A comparison of the structure of N1-meA:T lesion with N6-meA:T base pair. (A) Side view structure of N1-meA10:T5′ with neighboring base pairs aligned with N6-meA10:T5′. N1-meA10:T5′ with neighboring base pairs appears in red, and N6-meA10:T5′ with neighboring base pairs in green. The figure illustrates a significant dihedral between N1-meA10 and N6-meA10 when two bases are overlaid together. (B) Overlay of N1-meA10 and adjacent T9 (in red) with N6-meA10 and neighbor T9 (in green) as viewed from top. The purine ring of N1-meA10 is in syn conformation, and the N6-meA10 base is in anti-conformation.
major groove. This base has a relatively poor electron density indicating its high flexibility. Instead of completely flipping out, the N3-meC base stays partially intrahelical (Figure 8A), with a tilt angle around 18° (0.6° for B-form DNA) and a twist angle of 49° (36° for B-form DNA). The distances of P_T4...P_N 3-meC5 and P_N 3-meC5...P_T6 are stretched to 7.2 Å and 6.9 Å, respectively. In contrast, the complementary guanine of N 3-meC remains intrahelical, stabilized by the stacking to A11' and A9', its two flanking bases. Water-mediated hydrogen bonds of N3_G10...O1_A11' and O1'_G10'...N3_A9' also contribute to the stabilization of the guanine. In addition, another water-mediated interaction between G10' and T4, N1_G10' to water (2.8 Å) and water to N1_T4 (3.1 Å), might also help to ‘lock’ the unpaired G10' intrahelix. Although the neighboring base pair, T4:A11', is not apparently influenced by this lesion, a marked change occurred to the base pair T6:A9'. This base pair is disrupted such that T6 adopts a syn conformation to form a Wobble base pair with the opposite T with the purine ring of N1-meA adopting a syn conformation, which is consistent with the previous NMR result (38). This crystal structure also provides detailed parameters on the N1-meA:T pair and its impact on the local DNA duplex structure. Due to steric clashes caused by the N1-methylated adenine, we observed a 3 Å shift of N1-meA base into the major groove in the structure, as well as significant changes in the values of the opening, tilt and twist angles. A shortening of C1'-C1', RN9'-YN1 and RC8'-YC6 distances between N1-meA and the opposite T is observed, probably induced by the displacement of the N1-meA to the major groove and the movement of the complementary thymine toward the center of the helix. As a result of these conformational changes, the neighboring base pair G11:C4' shifts to the 3' direction, while the 5' neighboring base pair T9:A6' is left almost unaffected. A detailed investigation of ABH2–N1-meA also reveals that the protein residues, which normally establish contacts with the backbone in the middle of the lesion-containing strand and 5'-end of the complementary strand, lose most of the hydrogen-bonding interactions with the backbone of the complementary strand (Figure 1B). This observation suggests that: (i) these interactions to the complementary lesions are efficiently detected and repaired by the AlkB family proteins (24–31). To structurally characterize these methylation bases in duplex DNA, a disulfide cross-linked ABH2–dsDNA complex was employed as a host–guest-like system for DNA crystallization. To minimize potential influence from the bound protein, we placed the methylated bases away from the interface between the protein and DNA.

This protein–DNA complex, stabilized by covalent cross-linking between the protein and DNA, affords the first crystal structures of N1-meA and N3-meC in duplex DNA. We show that N1-meA forms a Hoogsteen base pair with the opposite T with the purine ring of N1-meA adopting a syn conformation, which is consistent with the previous NMR result (38). This crystal structure also provides detailed parameters on the N1-meA:T pair and its impact on the local DNA duplex structure. Due to steric clashes caused by the N1-methylated adenine, we observed a 3 Å shift of N1-meA base into the major groove in the structure, as well as significant changes in the values of the opening, tilt and twist angles. A shortening of C1'-C1', RN9'-YN1 and RC8'-YC6 distances between N1-meA and the opposite T is observed, probably induced by the displacement of the N1-meA to the major groove and the movement of the complementary thymine toward the center of the helix. As a result of these conformational changes, the neighboring base pair G11:C4' shifts to the 3' direction, while the 5' neighboring base pair T9:A6' is left almost unaffected. A detailed investigation of ABH2–N1-meA also reveals that the protein residues, which normally establish contacts with the backbone in the middle of the lesion-containing strand and 5'-end of the complementary strand, lose most of the hydrogen-bonding interactions with the backbone of the complementary strand (Figure 1B). This observation suggests that: (i) these interactions to the complementary

DISCUSSION

Cytotoxic damages such as N1-meA and N3-meC can block Watson–Crick base pairing in duplex DNA, which leads to cytotoxic consequences to living cells. These base

Figure 7. Diagram of the DNA structure containing N3-meC:G. (A) Structure of dsDNA2 with N3-meC:G. N3-meC5 in magenta, opposite G10' in blue. dsDNA 2: 5'-CTGT(3meC)TC*ATTGCG-3' paired with 5'-TCGCAATAAGACA-3'. (B) Structure of dsDNA2 with N3-meC:G aligned with the unlesioned dsDNA in structure 3BTX. DNA containing N3-meC in red, unlesioned DNA in green. dsDNA in 3BTX: 5'-CTGTATC*ATT GCG-3' paired with 5'-TCGCAATAAGACA-3'. Different packing pattern at the end of the dsDNA is observed.
strand might be weak, perhaps non-existent when the ABH2 protein is scanning through the duplex DNA to detect potential base lesions; (ii) the covalent cross-linking is important for holding the complex together, thus providing necessary robustness to this host–guest-like system with a low binding affinity in the absence of cross-linking.

We also report the structure of the ABH2–dsDNA complex containing the N³-meC lesion. The electron density of N³-meC is poor due to the absence of any hydrogen-bonding interaction to the opposing G. However, N³-meC remains intrahelical with a movement of ~2 Å to the major groove. The opposite guanine is also intrahelical and stacks to the neighboring A11 and A9. The disruption of the N³-meC5:G10 base pair also yields an elongated back-bone P...P distance between C5 to T4 and T6. Furthermore, the flanking base pair between T6 and A9 is broken, which leads to a Wobble base pairing between T6 and A8. The melting of the DNA structure around the lesion site generates low thermodynamic stability and produces the distortion of the DNA backbone, an effect that may lead to the loss of interactions between the protein and the backbone around the A6–C4 stretch on the complementary strand. These changes in DNA structure may explain why this lesioned base cannot be crystallized through typical DNA-crystallization strategies.

Both N¹-meA and N³-meC are efficiently detected and repaired by the AlkB family proteins, yet they show distinct structural features in duplex DNA. The only common property of the base pairs containing these two lesions is the low thermodynamic stability of the local duplex DNA structure induced by the methylation.

Thus, the repair proteins could detect unstable regions in the genome to locate these base damages.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the staff at beamline 23BM (General Medicine and Cancer Institute Collaborative Access Team [GM/CA-CAT]) at the Advanced Photon Source at Argonne National Laboratory for data collection, and Dr E. Duguid for data processing.

FUNDING

National Institutes of Health (GM071440 to C.H.); Beamline 23ID-B (General Medicine and Cancer Institutes Collaborative Access Team ([GM/CA-CAT]) at the Advanced Photon Source at Argonne National Laboratory; National Institutes of Health and the United States Department of Energy. Funding for open access charge: GM071440.

Conflict of interest statement. None declared.

REFERENCES

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. Nature, 362, 709–715.
2. De Bont, R. and van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, 19, 169–185.

3. Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, A.R. and Ellenberger, T. (2005) Biological response to DNA damage. In Friedberg, E.C. (ed.), *DNA Repair and Mutagenesis*. ASM Press, Washington DC, Chapter 1, pp. 2–7.

4. Lindahl, T. (2004) Molecular biology: ensuring error-free repair. *Nature*, 427, 598.

5. Wood, R.D., Mitchell, M., Sguojos, J. and Lindahl, T. (2001) Human DNA repair genes. *Science*, 291, 1284–1289.

6. Truglio, J.J., Croteau, D.L., Van Houten, B. and Kisher, C. (2006) Prokaryotic nucleotide excision repair: the UvRABC system. *Chem. Rev.*, 106, 233–252.

7. Sancar, A. (1995) DNA repair in humans. *Annu. Rev. Genetics*, 29, 69–105.

8. Egli, M. (2004) Nucleic acid crystallography: current progress. *Curr. Opin. Chem. Biol.*, 8, 580–591.

9. Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. USA*, 78, 2179–2183.

10. Hossain, M.T., Sunami, T., Tsunoda, M., Hikima, T., Chatake, T., Hikima, T. and Georgiadis, M.M. (2000) Crystal structure of an N-terminal fragment of the RNA-binding protein U1A as a crystallization module. *Nature*, 407, 567–574.

11. Meckler, G. (2000) Structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein UIA as a crystallization module. *J. Mol. Biol.*, 295, 1–16.

12. Ye, J.-D., Terseshkov, V., Frederiksen, J.K., Koide, A., Fellouse, F.A., Sidhu, S.S., Kide, S., Kossiakoff, A.A. and Piccirilli, J.A. (2000) Synthetic antibodies for specific recognition and crystallization of structured RNA. *Proc. Natl Acad. Sci. USA*, 105, 82–87.

13. Xiao, H., Morikami, H., Suga, H. and Fefer-D’Amare, A.R. (2008) Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *Nature*, 454, 358–361.

14. Okada, C., Yamashita, E., Lee, S.J., Shibata, S., Katakura, J., Nakagawa, A., Yoneda, Y. and Tsukihara, T. (2009) A high-resolution structure of the pre-microRNA nuclear export machinery. *Science*, 326, 1275–1279.

15. Lee, S., Bowman, B.R., Ueno, Y., Wang, S. and Verdine, G.L. (2008) Repair of alkylated DNA: recent advances. *DNA Repair*, 7, 429–442.

16. Tewrcw, S.C., Henschel, T., Hausinger, R.P., Lindahl, T. and Sedgwic, B. (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature*, 419, 174–178.

17. Falnes, P.O., Johansen, R.F. and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. *Nature*, 419, 178–182.

18. Delaney, J.C. and Essigmann, J.M. (2004) Mutagenesis, genotoxicity, and repair of 1-methyadenine, 3-alkycytosine, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proc. Natl Acad. Sci. USA*, 101, 14051–14056.

19. Mishina, Y., Yang, C.-G. and He, C. (2005) Direct repair of the exocyclic DNA adduct 1,6-ethenoadenine by the DNA repair AlkB proteins. *J. Am. Chem. Soc.*, 127, 14594–14595.

20. Dinley, J.C., Smeerger, L., Wong, C., Frick, L.E., Taghizadeh, K., Wishnok, J.S., Drumman, C.L., Samson, I.-D. and Essigmann, J.M. (2005) AlkB reverses etheno DNA lesion caused by lipid oxidation in vitro and in vivo. *Nat. Struct. Mol. Biol.*, 12, 855–860.

21. Aas, P.A., Otterlei, M., Falnes, P.O., Vågbo, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjorås, M., Sluppheng, G., Seeberg, E. et al. (2005) Human and bacterial oxidative demethylases repair alkyladenine damage in both RNA and DNA. *Nature*, 421, 859–863.

22. Westbye, M.P., Feyz, E., Aas, P.A., Vågbo, C.B., Talstad, V.A., Kavli, B., Hagen, L., Sundheim, O., Akbari, M., Llabakk, N.-B. et al. (2008) Human AlkB homolog I is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. *J. Biol. Chem.*, 283, 25046–25056.

23. Ougland, R., Hang, C.-M., Lilv, A., Johansen, R.F., Seeberg, E., Hou, Y.-M., Remme, J. and Falnes, P.O. (2004) AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Mol. Cell.*, 16, 107–116.

24. Duncan, T., Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T. and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proc. Natl Acad. Sci. USA*, 99, 16660–16665.

25. Ringvoll, J., Nordstrand, L.M., Vågbo, C.B., Talstad, V., Reite, K., Aas, P.A., Lauritzen, K.H., Liabakk, N.B., Bjørk, A., Doughty, R.W. et al. (2006) Repair deficient mice reveal mABH2 as the primary oxidative demethylation for repairing 1mCm and 3mCm lesions in DNA. *EMBO J.*, 25, 2189–2198.

26. Yang, C.-G., Yi, C., Duguid, E.M., Sullivan, C.T., Jian, X., Rice, P.A. and He, C. (2008) Crystal structure of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature*, 452, 961–965.

27. Read, R.J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. D.*, 57, 1373–1382.

28. Jones, J.W. and Robins, R.K. (1963) Purine nucleosides. III. Methylation studies of certain naturally occurring purine nucleosides. *J. Am. Chem. Soc.*, 85, 193–201.

29. Lu, X.-J. and Olson, W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, 31, 5108–5121.

30. Yang, H., Zhan, Y., Feng, D., Chi, L.M. and Lam, S.L. (2008) Effect of 1-methyadenine on double-helical DNA structures. *FEBS Lett.*, 582, 1629–1633.

31. Aishima, J., Gitti, R.K., Noah, J.E., Gan, H.H., Schluck, T. and Wolberger, C. (2002) A highosteen base pair embedded in undistorted B-DNA. *Nucleic Acids Res.*, 30, 5244–5252.

32. Abresch, N.G., Thoma, H., Huyhn-Dinh, T. and Siribara, J.A. (2002) Crystal structure of an antiparallel DNA fragment with highosteen base pairing. *Proc. Natl Acad. Sci. USA*, 99, 2806–2811.