Crystallographic Studies on Damaged DNAs. II. \(N^6\)-Methoxyadenine can Present Two Alternate Faces for Watson-Crick Base-pairing, Leading to Pyrimidine Transition Mutagenesis

Toshiyuki Chatake\(^1\), Takaaki Hikima\(^1\), Akira Ono\(^2\), Yoshihito Ueno\(^3\) Akira Matsuda\(^3\) and Akio Takénaka\(^1\)*

\(^1\)Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuda, Midori-ku, Yokohama, 226-8501, Japan

\(^2\)Graduate School of Science Tokyo Metropolitan University Hachioji, Minamiosawa, Tokyo 192-0364, Japan

\(^3\)Graduate School of Pharmaceutical Sciences Hokkaido University, Nishi 6 Kita 12, Kita-ku, Sapporo 060-0812, Japan

In a previous paper, 2'-deoxy-\(N^6\)-methoxyadenosine (mo\(^6\)A) was shown to form a mismatch base-pair with 2'-deoxycytidine with a Watson-Crick-type geometry. To fully understand the structural basis of genetic mutations with damaged DNA, it is necessary to examine whether the methoxylated adenine residue still has the ability to form the regular Watson-Crick pairing with a thymine residue. Therefore, a DNA dodecamer with the sequence d(CGCGmo\(^6\)AATTCGCG) has been synthesized and its crystal structure determined. The methylation has no significant effect on the overall DNA conformation, which is that of a standard B-form duplex. The methoxylated adenine moieties adopt the amino tautomer with an anti conformation around the C\(^6\)-N\(^6\) bond to the N\(^1\) atom, and they form a Watson-Crick base-pair with thymine residues on the opposite strand, similar to an unmodified adenine residue. It is concluded that methoxylated adenine can present two alternate faces for base-pairing, thanks to the amino \(\rightarrow\) imino tautomerism allowed by methylation. Based on this property, two gene transition routes are proposed.

© 1999 Academic Press

**Keywords:** \(N^6\)-methoxyadenine; mutagenesis; X-ray structure; mispairing; damaged DNA

**Introduction**

Genetic information is transmitted from generation to generation through DNA replication, in which a new complementary strand is synthesized starting from either strand of an original DNA duplex. In this process, the highest accuracy is achieved by forming Watson-Crick base-pairs between adenine and thymine bases and between guanine and cytosine bases, as an absolute rule in every organism. When DNA is attacked with certain chemicals, however, this rule is disturbed and errors are introduced into the synthesized DNA. As a result, genetic mutations will occur. These chemicals are called mutagens (Singer & Kusmierek, 1982). As mutations can lead to several diseases, including cancer, extensive studies on mutagens have been performed.

Oxyamines such as hydroxylamine and methoxylamine are known to be mutagens (Singer & Kusmierek, 1982) which predominantly attack and modify the exocyclic amino groups of nucleic acid bases (Kochetkov & Budowsky, 1969). Substitution at the N\(^6\) amino groups of adenine bases has been established (Budowsky *et al*., 1975), and is thought to be the origin of the observed mutations. It is expected that the N\(^6\)-methoxyadenine moiety has different chemical and physicochemical properties from the unmodified base. When a DNA fragment containing 2'-deoxy-N\(^6\)-methoxyadenosine (hereafter designated as mo\(^6\)A) is used as a template, the non-complementary 2'-deoxycytidine 5'-triphosphate (dCTP) as well as the complementary thymidine 5'-triphosphate (TTP) are incorporated by
the Klenow fragment of DNA polymerase I into the corresponding sites of the newly synthesized DNA strand (Nishio et al., 1992). In addition, free 2'-deoxyadenosine 5'-triphosphate (dATP), which is a reactant to be incorporated in the synthesized DNA, is also modified by oxyamines (Singer, 1975). Methoxylated dATP (hereafter designated as mo6dATP) is incorporated at both templates, 2'-deoxycytidine and thymidine (Singer & Spengler, 1982; Abdul-Maish & Bessman, 1986; Hill et al., 1998). The DNA polymerase accepts only Watson-Crick base-pairs, therefore it is plausible that methoxylation allows the adenine moiety to form Watson-Crick-type pairs with both thymine and cytosine. Indeed, we recently determined the crystal structure of a DNA dodecamer with the nucleotide sequence d(CGCGmo6AATCCGCG) (hereafter designated as Dmo6A-C), and found that the mo6A residue forms a Watson-Crick-type pair with cytosine residues in B-form duplex DNA (Chatake et al., 1999). The adenine moiety has the imino tautomeric form, with the methoxy group in the anti conformation around the C6-N6 bond to the N1 atom. X-ray analyses of the related derivatives N9-benzyl-N6-methoxyadenine (Fujii et al., 1990) and N6-methoxy-2',3',5'-tri-O-methyladeno-

Figure 1 (legend opposite)
sine (Birnbaum et al., 1984) also indicate that the \( N^6 \)-methoxylated base takes the imino form, although with the methoxy group in the syn conformation. In solution, however, these derivatives are in equilibrium between the imino form and the amino form (Stolarski et al., 1984), the latter being preferred by the unmodified adenine base (Wolfenden, 1969). To understand the structural basis of the mutation mechanism, it is necessary to confirm that the methoxylated adenine residue can adopt the amino tautomer, and that it still has the ability to form a canonical Watson-Crick pair with thymine residues. For this purpose, we have synthesized a DNA dodecamer with the sequence \( d(CGCGmo6AATTCGCG) \) (Dmo6A·T) and determined its crystal structure. Electron density maps clearly show that the mo6A residue does indeed form Watson-Crick-type pairs with thymine residues in the \( B \)-form duplex. This demonstrates the existence of two alternate faces of the methoxylated 2'-deoxyadenosine through the amino \( \rightarrow \) imino tautomerism, allowing its base-pairing with either 2'-deoxycytidine or thymidine, and indicates several possible gene transition mechanisms.

Results and Discussion

Crystal structures

The two Dmo6A·T strands are associated to form a right-handed double helix as shown in Figure 1(a). All torsion angles and local helical parameters (calculated by the program NUPARM; Bansal et al., 1995) are given in the Supplementary Material (Tables 1 and 2, respectively). The crystal structure is isomorphous to the original dodecamer with the sequence \( d(CGCGAATTCGCG) \) (Dickerson et al., 1981; Shui et al., 1998), and to the modified (“damaged”) dodecamer reported in the previous paper (Chatake et al., 1999). Similar to those structures, two duplexes related by \( 2_1 \) symmetry along the \( c \)-axis form columns in a head-to-tail fashion. To connect them, two inter-duplex base-pairs are formed through \( N^2-H \cdots N^3 \) hydrogen bonds between \( G_{a12} \) and \( G_{a2} \), and between \( G_{b2} \) and \( G_{b12} \) (see Figure 2 by Chatake et al., 1999). An octahedrally hydrated magnesium cation is located in the major groove of each duplex. Three of the bound water molecules form hydrogen bonds with the \( O^6 \) and \( N' \) atoms of \( G_{a2} \) and \( G_{b10} \) and the other three are hydrogen bonded to the phosphate oxygen atoms of \( A_{a6} \) and \( T_{a7} \) in another duplex, related by \( 2_1 \) symmetry along the \( b \)-axis. This situation is the same as that found in the Dmo6A·C crystal (see Figure 4 by Chatake et al., 1999).

Effect of methoxylation on DNA conformation

The structures of both Dmo6A·T (Figure 1(a)) and Dmo6A·C (Figure 1(b)) are similar to that of the original dodecamer (Figure 1(c)). The overall root-mean-square displacement of non-hydrogen atoms is 0.31 Å for Dmo6A·T and 0.31 Å for Dmo6A·C, when they are superimposed on the original dodecamer. Figure 2 shows plots of some representative helical parameters along the nucleotide sequence. These parameters fluctuate around average values close to those of typical \( B \)-form

![Figure 1](https://example.com/figure1.png)

Figure 1. Stereo views of DNA dodecamers (a) Dmo6A·T, (b) Dmo6A·C (Chatake et al., 1999) and (c) the original dodecamer (Dickerson, et al., 1981; Shui et al., 1998). The diagrams were drawn with the program MOLSCRIPT (Kraulis, 1991). The mo6A residues are shown in red and the 2'-deoxycytidine residues paired with mo6A are in green.
DNA, and their variations along the sequence are almost the same, even at the mo6A residues, indicating that methoxylation of adenine residues does not significantly affect the overall DNA conformation. The X-ray structure (Kiefer et al., 1998) of a large fragment of DNA polymerase I from Bacillus stearothermophilus, complexed with DNA primer and template strands, shows that the two DNA strands form a duplex with a B-form conformation in the polymerase active site. In addition, there is an open, solvent-accessible space in the major groove of the bound DNA. Therefore, when adenine bases in DNA are methoxylated, the damaged DNA can still adopt an acceptable conformation in the polymerase, and the methoxy group will not interfere with the binding.

Hydrogen-bonding scheme of mo6A

Figure 3 shows 2|Fo| − |Fc| maps at the two sites of the mo6A residues. Each N6-methoxyadenine moiety forms a pair with the thymine base on the opposite strand. It is noteworthy that the geometry of the pairing is essentially the same as the usual Watson-Crick adenine-thymine pair. The methoxy groups of the mo6A bases take an anti conformation with respect to the N1 atom around the C6-N6 bond, like those found in the Dmo6A-C crystal (Chatake et al., 1999). The hydrogen bond distances and angles in the base-pairs are given in Table 1: those of the other hydrogen bonds are in the Supplementary Material (Table 3). The distances between the N6 (mo6Aa5) and O4 (Tb8) atoms and between the N1 (mo6Aa5) and N3 (Tb8) atoms, and the corresponding distances in the mo6Ab5-Ta8 pair, indicate that mo6A can form a base-pair through hydrogen bonds with thymine just like the unmodified adenine residue. In the Dmo6A-C crystal, a Watson-Crick-like mo6A-C pairing was

Figure 2. A comparison of some representative helical parameters in Dmo6A-T (triangles), Dmo6A-C (Chatake et al., 1999) (circles) and the original dodecamer (Dickerson et al., 1981; Shui et al., 1998) (boxes). The top, middle and bottom diagrams represent the helical rise, the helical twist angles and the displacements, respectively, of the three duplexes. They are plotted along the nucleotide sequence to show their fluctuations. These parameters were calculated by the program NUPARM (Bansal et al., 1995). In the right shaded column, their average values are indicated, together with the typical ones for standard A and B-form DNA (Arnot & Hukins, 1972).

Figure 3. Final 2|Fo| − |Fc| electron density on the 2′-deoxy-N6-methoxyadenosine-thymidine pairs mo6Aa5-Tb8 and mo6Ab5-Ta8 in Dmo6A-T. Broken lines show possible hydrogen bonds. Maps were contoured at the 1.4σ level by the program O (Jones et al., 1991). The nucleotides are numbered from the 5′ end independently in the two strands, a and b.
found (Chatake et al., 1999). When both the mo6A- T pair and the mo6A- C pair are superimposed on the canonical A- T pair found in the original dodecamer (Dickerson et al., 1981; Shul et al., 1998), the base positions are almost the same. In other words, the methoxylated adenine base can form a pair with either thymine or cytosine bases in the same geometry as a Watson-Crick pair, with no significant changes in atomic positions. Therefore, these pairings can be accepted as Watson-Crick pairs in the polymerase (Kiefer et al., 1998).

**Tautomerism of mo6A and base-pairing**

In the crystalline state, only the imino form has been found in N⁶-methoxyadenine derivatives (Fujii et al., 1990; Birnbaum et al., 1984). The imino form is also preferred in the Dmo6A-C crystal, forming a Watson-Crick-type base-pair with a cytosine residue (Chatake et al., 1999). In the present Dmo6A-T crystal, however, to form the observed base-pairs with thymine residues, the chemical structure of the methoxylated adenine base must be the amino form, as shown in Figure 4. An unmodified adenine base almost always exists in the amino form, its tautomerization occurring very rarely (Wolfenden, 1969). The existence of either tautomer of mo6A in the crystalline state, depending on the base in the opposite strand, is consistent with the fact that in solution, an equilibrium is observed between the amino and the imino forms (Stolarski et al., 1984), the ratio between the two tautomers being a function of the polarity of the solution. Furthermore, when a cytosine derivative is added to a solution containing an mo6A derivative, the imino form increases in population, and when a uridine derivative is added, the amino form increases (Stolarski et al., 1987). Therefore, it is concluded that it is the formation of a Watson-Crick-type base-pair with either 2'-deoxy- thymidine or thymidine that stabilizes one or the other tautomer.

The difference in total energy between the two tautomers and the stabilization energy on base-pairing were calculated by the molecular orbital method (see Figure 5). When 9-methyl-N⁶-methoxyadenine is alone in vacuum, the amino form is more stable than the imino form, as in a hydrophobic solution. But when mo6A makes a base-pair with thymine or with cytosine in the Watson-Crick type, the resulting base-pairs are stabilized at almost the same level, consistent with the results of the present X-ray analyses that 2'-deoxy-N⁶-methoxyadenosine can form a base-pair with either thymidine or 2'-deoxythymidine.

In summary, when a thymine residue is located on the opposite site in a DNA duplex, mo6A takes the amino tautomer and behaves like an unmodified adenine, forming a base-pair with the thymidine. On the other hand, when a cytosine residue is located on the opposite strand, as discussed in the previous paper for Dmo6A-C, the mo6A residue takes the imino form and behaves like a 2'-deoxyguanosine, forming a base-pair with the two hydrogen bonds between N⁴ (mo6A) and N⁴ (C) and between N⁶ (mo6A) and N⁴ (C). The existence of such “alternate faces” of 2'-deoxy-N⁶-methoxyadenosine in base-pairing points clearly to a possible mutation mechanism, discussed below.

**Biological implication**

The possibility of two alternate faces of N⁶-methoxyadenine is applicable to adenine residues in template DNA strands and to dATP as a reactant being incorporated during DNA replication. Considering with successive replication steps, possible routes of gene transition are shown in Figure 6.

![Figure 4. Chemical structures for mo6A-T and mo6A-C pairings.](image-url)
When an adenine base in DNA is methoxylated, the original A-T pair can be replaced by a G-C pair after two steps. In the first step, if dCTP is incorporated instead of TTP at the methoxylated adenine residue, then dGTP can be incorporated at the 2'-deoxycytidine site in the second step. In the case of mo6dATP incorporation, A-T to G-C and G-C to A-T mutations are possible with three replication steps. In the A-T → G-C case, if mo6dATP is incorporated at the template thymine residue by forming a mo6A-T pair, then the incorporated mo6A residue will next accept dCTP by mimicking G; finally, dGTP will be incorporated at the cytosine residue of the second template. In the G-C → A-T case, if mo6dATP is incorporated at the template cytosine residue by mimicking G, then the incorporated mo6A residue will accept TTP; finally dATP will be incorporated at the template thymine residue.

**Experimental Procedures**

**Synthesis and crystallization**

DNA dodecamer of Dmo6A-T was synthesized and crystallized in the same way as that reported in the previous paper (Chatake et al., 1999). A reservoir solution containing 25% (v/v) 2-methyl-2,4-pentanediol (MPD), 18 mM magnesium acetate, 6 mM spermine tetrahydrochloride, 80 mM sodium chloride and 10 mM sodium cacodylate at pH 7.0 was used.
Data collection

To prevent nucleation and growth of hexagonal ice, the crystals were bathed in a reservoir solution containing 35% (v/v) MPD for one minute before flash freezing. X-ray data were collected at 110 K on the Sakabe-Weissenberg camera (Sakabe, 1991) with synchrotron radiation at the Photon Factory (BL-6B) in Tsukuba. To compensate the blind region, another crystal with a different orientation was tested at 100 K. The two sets of diffraction patterns were processed by the program DENZO (Otwinowski & Minor, 1997), and intensity data of a total of 28,875 observed reflections were merged by the programs SCALA and AGROVATA in CCP4 (Collaborative Computational Project Number 4, 1994) into 7518 independent reflections with R_merge of 3.6%. The completeness of the data was 79.7% in the 100 ~ 1.6 Å resolution range and 59.5% for the outer 1.67 ~ 1.60 Å resolution shell. The unit cell dimensions are \( a = 25.5 \text{ Å}, b = 39.8 \text{ Å} \) and \( c = 66.5 \text{ Å} \), the space group being \( P2_12_12_1 \). Statistics of data collection and crystal data are given in Table 2.

Structure determination

The initial structure was determined by the molecular replacement method, using the structure of the original DNA dodecamer \( d(CGCGAATTCGCG) \) (Shui et al., 1998) with the program AMoRe (Navaza, 1994). To inspect the molecular structure and to add water molecules, the programs O (Jones et al., 1991) and QUANTA (distributed by Molecular Simulations, Inc.) were used. Taking the hydrogen-bonding scheme into consideration, the moA bases were assumed to have an amino form. For X-PLOR (Brünger, 1992a) refinement, the stereochemical parameters of the amino form of moA were derived by combination of the structures of an unmodified 2-deoxy-adenosine and the methoxy group of 11-methoxygelseamide (Lin et al., 1989). A combination of simulated annealing and positional refinements was applied, followed by interpretation of omit maps at every nucleotide residue. One magnesium cation hydrated octahedrally and one spermine molecule was found, and 126 water molecules were assigned. During structural refinement, no constraints were applied between the paired moA and thymidine residues. The final R-factor was 21.1% for 5.0 ~ 1.6 Å resolution data (R_free = 27.0% for 10% of the observed data). Statistics of the structure determination are summarized in Table 2.

Table 2. Crystal data, data collection and structure determination

<table>
<thead>
<tr>
<th>Space group</th>
<th>( P2_12_12_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit cell (Å)</td>
<td>( a = 25.5, b = 39.8, c = 66.5 )</td>
</tr>
<tr>
<td>Asymmetric unit (duplexes)</td>
<td>4</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>100 ~ 1.6</td>
</tr>
<tr>
<td>Measured reflections</td>
<td>28,875</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>7518</td>
</tr>
<tr>
<td>Completeness (%) in the outer shell (%)</td>
<td>59.5(1.67 ~ 1.60 Å)</td>
</tr>
<tr>
<td>( R_{merge} ) (%)</td>
<td>3.6</td>
</tr>
<tr>
<td>Structure refinement</td>
<td>Resolution range 5.0 ~ 1.6</td>
</tr>
<tr>
<td>Used reflections</td>
<td>6836</td>
</tr>
<tr>
<td>Final model</td>
<td>Number of DNA atoms 490</td>
</tr>
<tr>
<td></td>
<td>Number of water molecules 126</td>
</tr>
<tr>
<td></td>
<td>Number of spermine atoms 14</td>
</tr>
<tr>
<td></td>
<td>Number of magnesium atoms 1</td>
</tr>
<tr>
<td>Average B-factors (Å²)</td>
<td>DNA 16.4</td>
</tr>
<tr>
<td></td>
<td>Water molecules 33.1</td>
</tr>
<tr>
<td></td>
<td>Spermine atoms 56.6</td>
</tr>
<tr>
<td></td>
<td>Magnesium atom 13.8</td>
</tr>
<tr>
<td>R-factor (%)</td>
<td>21.1</td>
</tr>
<tr>
<td>( R_{free} ) (%)</td>
<td>27.0</td>
</tr>
<tr>
<td>r.m.s. deviation</td>
<td>Bond lengths (Å) 0.005</td>
</tr>
<tr>
<td></td>
<td>Bond angles (deg.) 1.1</td>
</tr>
<tr>
<td></td>
<td>Improper angles (deg.) 1.6</td>
</tr>
<tr>
<td>Average coordinates error (Å)</td>
<td>0.22</td>
</tr>
</tbody>
</table>

\( R_{merge} = 100 \times \frac{\sum_{hkl} |F_{o}| - |F_{c}|}{\sum_{hkl} |F_{o}|} \)
\( R \) factor = \( 100 \times \frac{\sum_{ij} |F_{o,ij}| - |F_{c,ij}|}{\sum_{ij} |F_{c,ij}|} \), where \( |F_{o,ij}| \) and \( |F_{c,ij}| \) are observed and calculated structure factor amplitudes, respectively.

\( R_{free} = \frac{\sum_{hkl} |F_{o,ij}| - |F_{c,ij}|}{\sum_{hkl} |F_{o,ij}|} \) calculated using a random set containing 10% of observations, that were omitted during refinement (Brünger, 1992b).

Data Bank accession code

The atomic coordinates have been deposited in the Nucleic Acid Database (NDB) (entry code BD0010).

Acknowledgments

We thank N. Sakabe and N. Watanabe for facilities and help during data collection at the Photon Factory (Tsukuba), and T. Simonson for proofreading of the manuscript. This work was supported in part by a grant for the RFTF(97L00503) (Research For The Future) from the Japanese Society for the Promotion of Science, by Grants-in-Aid for Scientific Research on Priority Area (Nos. 08124203 and 07250205) from the Ministry of Education, Science, Sport and Culture of Japan, and by the Sakabe project of TARA (Tsukuba Advanced Research Alliance), University of Tsukuba.

References


Edited by I. Tinoco