SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2010, Article ID 107289, 10 pages
doi:10.4061/2010/107289

Research Article

Insights into the Structures of DNA Damaged by Hydroxyl Radical: Crystal Structures of DNA Duplexes Containing 5-Formyluracil

Masaru Tsunoda,[1] Takeshi Sakaue,[2,3] Satoko Naito,[2,4] Tomoko Sunami,[2,5] Naoko Abe,[6,7] Yoshihito Ueno,[6,8] Akira Matsuda,[6] and Akio Takénaka[1,2,9]

[1] *Faculty of Pharmacy, Iwaki Meisei University, Chuodai-Iino, Iwaki 970-8551, Japan*
[2] *Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama 226-8501, Japan*
[3] *Teijin Institute for Biomedical Research, Teijin Ltd., 4-3-2 Asahigaoka, Hino, Tokyo 191-0053, Japan*
[4] *Discovery Technology Research Laboratories, Eisai Co. Ltd., 13 Tokodai 5-chome, Tsukuba-shi, Ibaraki 300-2635, Japan*
[5] *Kansai Photon Science Institute, JAEA, 8-1-7 Umemidai, Kizugawa 619-0215, Japan*
[6] *Graduate School of Pharmaceutical Sciences, Hokkaido University, Kita-ku, Sapporo 060-0812, Japan*
[7] *Nano Medical Engineering Laboratory, Advanced Science Institute, RIKEN 2-1, Hirosawa, Wako-Shi, Saitama 351-0198, Japan*
[8] *Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan*
[9] *Graduate School of Science and Technology, Iwaki Meisei University, Chuodai-Iino, Iwaki 970-8551, Japan*

Correspondence should be addressed to Akio Takénaka, atakenak@bio.titech.ac.jp

Received 7 June 2010; Revised 19 July 2010; Accepted 15 August 2010

Academic Editor: Ashis Basu

Hydroxyl radicals are potent mutagens that attack DNA to form various base and ribose derivatives. One of the major damaged thymine derivatives is 5-formyluracil (fU), which induces pyrimidine transition during replication. In order to establish the structural basis for such mutagenesis, the crystal structures of two kinds of DNA d(CGCGRATfUCGCG) with R = A/G have been determined by X-ray crystallography. The fU residues form a Watson-Crick-type pair with A and two types of pairs (wobble and reversed wobble) with G, the latter being a new type of base pair between ionized thymine base and guanine base. *In silico* structural modeling suggests that the DNA polymerase can accept the reversed wobble pair with G, as well as the Watson-Crick pair with A.

1. Introduction

Hydroxyl radicals, activated from hydrogen peroxide and hydrogen superoxide anion under light radiation, are well known as potent mutagens that attack DNA and convert them to many different kinds of base and ribose derivatives [1, 2]. Every aerobic organism possesses several enzymes to remove such toxic oxides, as well as to recover the damaged DNA. However, when an excess amount of the radicals attacks DNA, the thymine base is oxidized at the 5-methyl group to form 5-formyluracil base (hereafter 2′-deoxy-5-formyluridine residue is referred to as fU) as a major product. (The four characters, A, T, G, and C, represent the respective nucleotide residues in DNA sequence. The other abbreviations used are fU for 5-formyluracil or 2′-deoxy-5-formyluridine residue, dfUTP for 2′-deoxy-5-formyluridine 5′-triphosphate, HPLC for high pressure liquid chromatography, fUA for d(CGCG**A**AT**fU**CGCG), and fUG for d(CGCG**G**AT**fU**CGCG).) It was demonstrated that 2′-deoxy-5-formyluridine triphosphate (dfUTP) was incorporated against both A and G templates, possibly forming fU:A and fU:G base pairs during *in vitro* DNA replication [3, 4]. On the other hand, it was reported that dfUTP-induced pyrimidine transitions, G:C → A:T and A:T → G:C, as well as a gene transversion from G:C to T:A, could occur *in vivo* [5, 6]. These results suggest that fU can behave as C, A, and G in addition to its original property of T.

Table 1: Crystallization conditions.

Crystals	fUA[1]	fUA[2]	fUA[3]	fUA[4]	fUG[1]	fUG[2]	fUG[3]
Droplet							
Sodium cacodylate buffer solution (mM)	20	20	20	20	20[a]	20	20
pH	6.5	6.0	7.0	6.0	8.1	7.0	7.0
DNA (mM)	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Spermine 4HCl (mM)	6	6	6	6	6	6	6
Sodium chloride (mM)	6	40	40	6	—	40	40
Potassium chloride(mM)	60	—	—	40	—	—	—
Magnesium chloride(mM)	6	10	—	—	—	10	10
Barium chloride(mM)	—	—	10	—	—	—	—
MPD (% in v/v)	5	5	5	5	5	5	5
Hoechst 33258 (mM)	—	—	—	—	—	0.5	—
DAPI (mM)	—	—	—	—	—	—	0.5
Reservoir solution							
MPD (%)	40	40	40	40	40	35	40
Temperature (K)	277	277	277	277	277	277	277

MPD: 2-methyl-2,4- pentanediol.
Hoechst 33258: 2′-(4-hydroxyphenyl)-6-(4- methyl-1-piperazinyl)-2,6′-bi-*1H*-benzimidazole.
DAPI: 4′,6-diamidino-2-phenylindole.
[a]Buffer solution of fUG[1] is 3-[4-(2-Hydroxyethyl)-1-piperazunyl]propanesulfonic acid.

In order to reveal the interaction geometry of the modified base fU, we performed X-ray analyses on fU-containing DNA duplexes. The fU residues were introduced into the self-complementary Dickerson-Drew-type dodecamer sequence, which is expected to be easy to crystallize. The DNAs used in this study are d(CGCG**A**AT**fU**CGCG) and d(CGCG**G**AT**fU**CGCG) and will be referred to as fUA and fUG, respectively. The fU base faces either an adenine or a guanine at the two sites in each duplex. Four fUA crystals and three fUG crystals were obtained under different conditions. Their crystal structures have been successfully determined at resolutions ranging from 1.5 to 3.0 Å. In the preliminary papers [7, 8], magnesium ion effects on crystallizations of fUA were discussed, but the detailed structure of base-pair formations and its biological significance were not published. In this paper we describe the structures of the base pairs formed between fU and G and between fU and A, based on which the pyrimidine transition induced by the oxidized thymine base will be discussed.

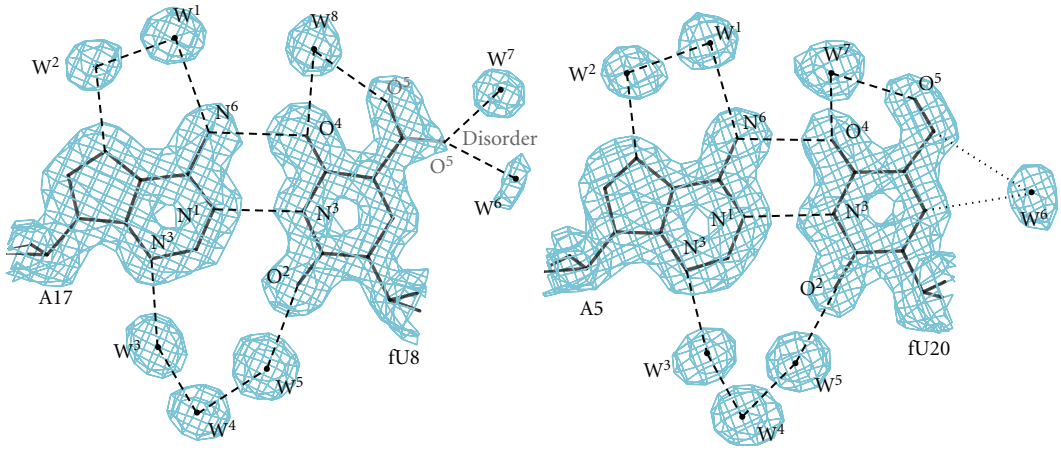
2. Materials and Methods

2.1. Oligodeoxyribonucleotide Synthesis. fUA and fUG were synthesized by the solid phase phosphoramidite method as described previously in [9] and were purified by reverse-phase column chromatography and reverse-phase and anion-exchange HPLCs. After $NaIO_4$ treatment, the oligonucleotides with fU were further purified by reverse-phase and anion-exchange HPLCs.

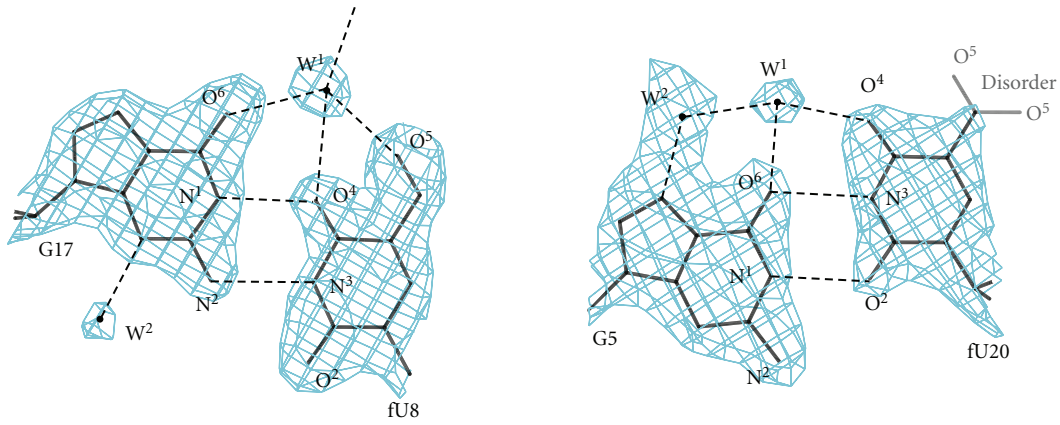
2.2. Crystallization and Data Collection. Initial screenings of crystallization conditions were performed using the hanging drop vapor diffusion method, equilibrating 2 μl droplets against 1 ml of the reservoir solution. The optimized conditions for growing the four different crystals of fUA (fUA[1], fUA[2], fUA[3], and fUA[4]) and three different crystals of fUG are given in Table 1. As the fUG[1] crystal was too small, two kinds of dyes, Hoechst 33258 and DAPI, were added to stabilize the duplex formation (fUG[2] and fUG[3]).

Crystals suitable for X-ray data collections were picked up from their mother liquors using a nylon loop (Hampton Research) and transferred into liquid nitrogen. All X-ray experiments for the seven crystals were performed with synchrotron radiation at the Photon Factory in Tsukuba. Diffraction patterns of the fUA crystals, which were recorded on imaging plates, were processed subsequently using the programs *DENZO* and *Scalepack* [10] and those of the fUG crystals, recorded on Quantum 4 CCD, with the program *DPS/MOSFLM* [11]. Low-resolution data around the 10 Å resolution shell were truncated because they did not fit well into the diffraction profile. The crystal data and the statistics of data collection are summarized in Table 2.

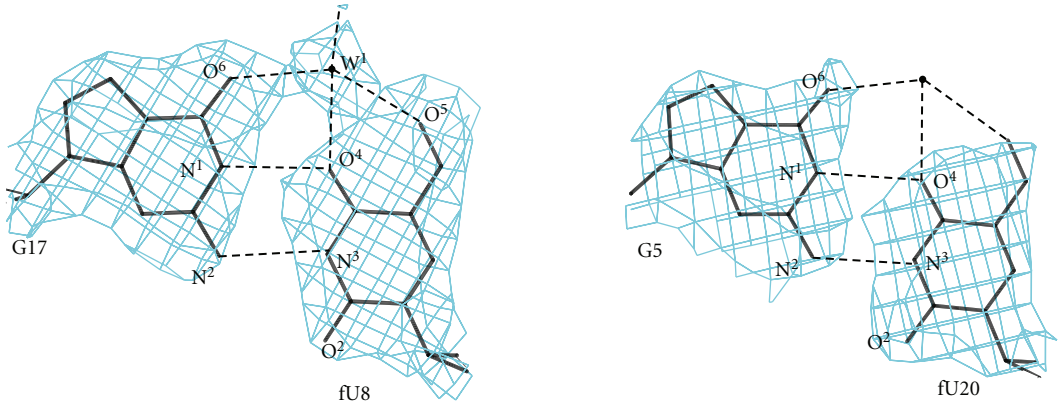
2.3. Structure Determination and Refinement. Initial phases were derived by molecular replacement with the program *AMoRe* [12] using the atomic coordinates of the corresponding unmodified DNA duplex (PDB ID 355D, see [13]) as structural probes. The molecular structures were constructed and modified on a graphic workstation with the program *Coot* [14] in the CCP4 program suite [15]. The atomic parameters were refined with the program *Refmac* [16] in CCP4 [15] with maximum-likelihood techniques, followed by interpretation of an omit map at every nucleotide residue. Newly defined patches were prepared for partially applying



(a)



(b)



(c)

FIGURE 1: The final $2|F_o|-|F_c|$ maps around the fU residues in the fUA[2] (a), the fUG[2] (b), and the fUG[3] (c) crystals. The maps are contoured at 1σ level by the program *DINO* [18]. Those of the fUG[1] crystal are omitted due to similarity to (c). Values indicate possible hydrogen bond distances. Characters N, O, and W indicate nitrogen, oxygen, and water oxygen atoms, respectively.

Table 2: Crystal data and statistical information on data collection and processing and on structure determination.

Crystals	fUA[1]	fUA[2]	fUA[3]	fUA[4]	fUG[1]	fUG[2]	fUG[3]
Wavelength (Å)	1.00	0.90	1.00	1.00	1.00	1.00	1.00
Beamline (at PF)	BL-18B	BL-6B	BL-18B	BL-6B	BL-18B	BL-18B	BL-18B
Oscillation ranges(°)	5	3	5	5	3	2	2
Frames	72	60	34	36	60	90×2^{c}	90
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a (Å)	25.8	26.0	25.2	25.3	24.6	25.3	25.0
b (Å)	39.3	39.5	41.2	41.7	40.0	40.3	40.5
c (Å)	65.0	65.7	65.4	66.0	68.5	66.0	66.8
Z^{a}	1	1	1	1	1	1	1
Matthews coefficient (Å^3 Da^{-1})	2.15	2.20	2.22	2.28	2.28	2.28	2.29
Solvent content (%)	42.9	44.2	44.6	45.9	46.0	46.1	46.3
Resolution (Å)	1.57	1.5	1.7	1.8	3.0	1.95	2.6
Completeness (%)	97.1	99.4	97.7	99.3	99.3	99.9	99.9
R^{b}_{merge} (%)	3.1	4.8	2.9	2.8	11.4	4.0	6.3
Observed reflections	87588	66227	31872	45728	16654	72307	38399
Unique reflections	9502	10301	7779	6881	1531	5312	2342
Redundancy	9.22	6.43	4.10	6.65	10.9	13.6	16.4
Structure refinement[f]							
Resolution range (Å)	10.0–1.57	10.0–1.55	10.0–1.85	10.0–1.80	15.0–3.02	10.0–1.95	10.0–2.70
R-factor (%)[d]	18.3	18.3	21.6	19.3	18.8	20.3	21.1
R_{free} (%)[e]	23.2	22.5	26.0	23.1	22.9	26.1	27.8
R.m.s. deviation							
Bond distances (Å)	0.027	0.028	0.019	0.021	0.005	0.019	0.007
Bond angles (°)	3.8	3.6	3.2	3.4	0.9	3.1	1.0
No. of ions	$1Mg^{2+}$, $1K^{+}$	$1Mg^{2+}$	—	—	—	$1Mg^{2+}$	$1Mg^{2+}$
No. of additive molecules	—	—	—	—	—	1Hoechst33258	1DAPI
No. of water molecules	150	185	82	142	27	123	45
PDB-ID	1G75	1G8N	1G8U	1G8V	3AJJ	3AJK	3AJL

[a]Number of duplexes in the asymmetric unit. [b]$R_{merge} = 100 \times \sum_{\mathbf{h}j} |I_{\mathbf{h}j} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}j} I_{\mathbf{h}j}$, where $I_{\mathbf{h}j}$ is the jth measurement of the intensity of reflection $\mathbf{h}$ and $\langle I_{\mathbf{h}} \rangle$ is its mean value. [c]Two data sets were taken for a crystal by changing exposure time to compensate overloaded reflections. [d]R-factor $= 100 \times \sum ||F_o| - |F_c|| / \sum |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively. [e]Calculated using a random set containing 10% of observations that were not included throughout refinement [17].

structural restrains to the modified residue. Water, ion, and dye molecules were assigned and included in the refinements. The program *CNS* [17] was used in the final refinements of fUG[1] and fUG[3] to stabilize the base pairs containing X with hydrogen bonds. The statistics of structure refinements are summarized in Table 2. Examples of the quality of the final electron density maps are depicted in Figure 1. Helical and local base-pair parameters [19], as well as the torsion angles and pseudorotation phase angles of sugar rings [19], were calculated using the program *3DNA* [20]. Some of them are shown in Table 3 and Figure 4.

2.4. Coordinates. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 1G75, 1G8N, 1G8U, 1G8V, 3AJJ, 3AJK, and 3AJL for fUA[1], fUA[2], fUA[3], fUA[4], fUG[1], fUG[2], and fUG[3], respectively.

3. Results

3.1. Quality of X-ray Analyses. All the crystals are isomorphous to the orthorhombic form of the unmodified duplex crystal. Crystallization of fUA was easy as expected, and the crystals obtained diffracted within the 1.5–1.8 Å resolution range. In every crystal, the formyl group of the fU8 residue on one of the two chains was disordered between the two alternative conformers, *anti* and *syn*, while that in the other chain adopted only the *syn* conformation. Relative occupancies of the disordered oxygen atoms, estimated

Table 3: A comparison of the local base-pair parameters at the modified pairs, calculated with the program *3DNA* [20].

	Pair	fUA^1	fUA^2	fUA^3	fUA^4	fUG^1	fUG^2	fUG^3	Ave	Unm
Shear (Å)	fU8:R	0.01	−0.03	−0.05	−0.08	−2.00	−2.24	−2.19	−0.94	−0.11
	R:fU20	0.05	0.08	−0.04	−0.01	1.90	−2.08	2.15	0.29	−0.04
Stretch (Å)	fU8:R	−0.19	−0.13	−0.08	−0.10	0.15	0.00	0.08	−0.04	−0.12
	R:fU20	−0.10	−0.11	−0.16	−0.16	0.13	−0.58	0.10	−0.13	−0.11
Stagger (Å)	fU8:R	0.10	0.11	−0.06	0.08	−0.33	−0.08	0.03	−0.02	−0.00
	R:fU20	0.09	0.00	0.00	0.10	−0.06	0.10	−0.05	0.03	0.01
Buckle (°)	fU8:R	−6.04	−4.89	−2.68	−4.97	−5.62	−10.56	−9.96	−6.39	−1.56
	R:fU20	8.11	6.02	9.91	9.34	6.78	8.68	4.50	7.62	4.72
Propeller (°)	fU8:R	−12.4	−12.4	−12.5	−11.9	−9.1	−16.3	−13.3	−12.6	−16.4
	R:fU20	−14.3	−14.5	−15.2	−15.3	−18.8	−14.6	−13.9	−15.2	−15.3
$C^{1'} \cdots C^{1'}$	fU8:R	10.4	10.5	10.4	10.4	11.0	11.2	11.0	10.5(WC)	10.5
(Å)	R:fU20	10.5	10.6	10.3	10.4	11.1	10.6	11.1	11.1(rw)	10.5

R: a purine residue, A or G, Ave: average, Unm: unmodified duplex [14], WC: Watson-Crick type and rw: reversed wobble.

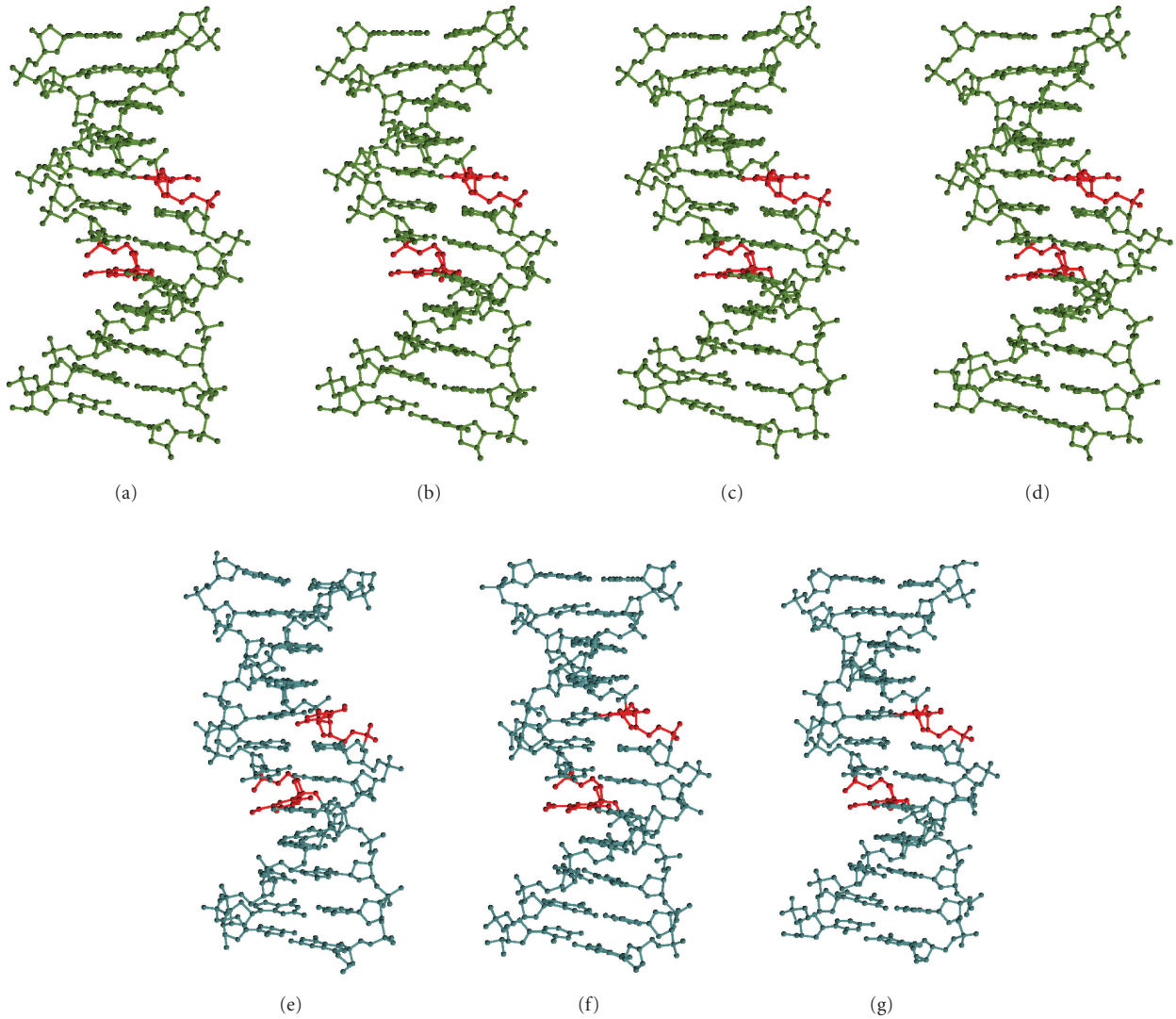


Figure 2: The overall structures of DNA duplexes containing fU, (a) fUA^1, (b) fUA^2, (c) fUA^3, (d) fUA^4, (e) fUG^1, (f) fUG^2, and (g) fUG^3. The fU residues are colored red. Hoechst33258 and DAPI are omitted for clarity.

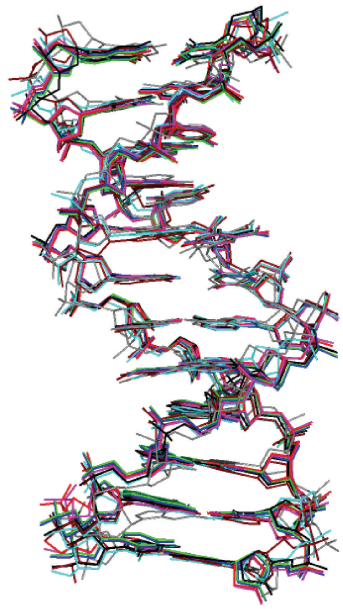


Figure 3: Superimposition of the fUA[1](green), fUA[2](blue), fUA[3](red), fUA[4](violet), fUG[1](gray), fUG[2](brown), and fUG[3](cyan) duplexes onto the unmodified duplex (black) for comparison of the duplex conformations.

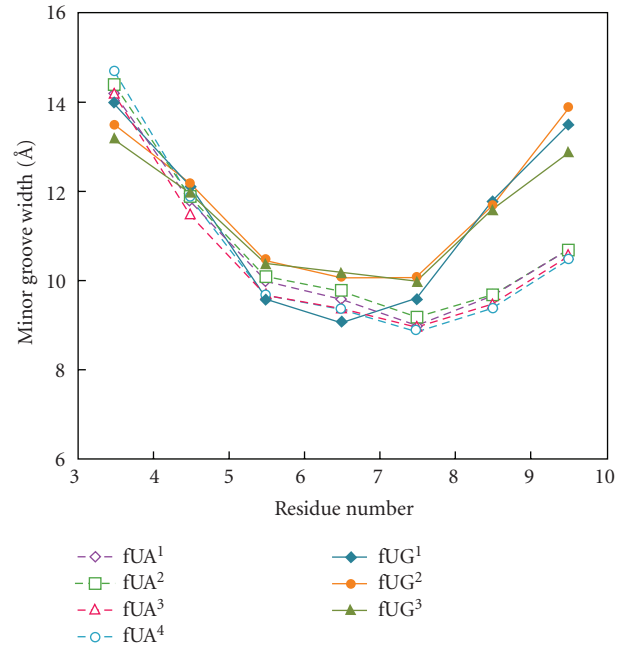


Figure 4: The minor groove widths in the fU-substituted and unmodified duplexes. The minor groove calculations were performed with the program *3DNA* [20]. The widths are wider at the central regions in fUG[2] and fUG[3] compared with those in fUA[1], fUA[2], fUA[3], fUA[4], and fUG[1] because Hoechst33258 and DAPI are bound in the grooves.

from the electron densities, were assumed in the structure refinements. (The residue numbering is 1 ~ 12 for one chain and 13 ~ 24 for another chain.)

On the other hand, fUG was difficult to crystallize. Attempts to crystallize fUG under conditions similar to those for fUA, that is, at neutral or slightly acidic pH, were unsuccessful. Small single crystals of fUG[2] appeared at basic pH (pH 8.1) but they poorly diffracted X-rays. Cocrystallization with several dyes was then attempted with the hope of stabilizing duplex formation. This approach led to the successful crystallization of fUG[2] and fUG[3] although the resolution of the fUG[3] crystal was still quite low. Hoechst 33258 and DAPI were found to be bound in the central region of the minor grooves of fUG[2] and fUG[3], respectively, in a manner similar to the other DNA duplexes containing such dyes [21–27]. Interestingly, during refinements of fUG[2], it was found that the two fU bases moved to different directions, one toward the minor groove side and the other toward the major groove side from the canonical Watson-Crick-type pairing position, and the resulting electron density also showed the same movements. In the case of fUG[1] and fUG[3], however, the two fU bases moved to the same direction toward the minor groove side, and the electron densities, though poor, also supported the movements.

3.2. Overall Structures. The local helical parameters show that all the fUA and the fUG duplexes adopt the B-form conformation even in complex with dyes, as shown in Figure 2. Superimpositions of the present structures onto the unmodified duplex structure are shown in Figure 3 and yield an average root mean square deviation of 1.4 Å. Closer inspection of the superimposed structures reveals no drastic differences between the modified and the unmodified duplexes. However, the minor groove widths (see Figure 4) indicate that fUG[2] and fUG[3] are wider at the center compared with those of the other DNA duplexes. These changes in the DNA conformation are presumably due to the binding of DAPI and Hoechst33258 rather than the fU introductions. Another feature is that, at residues 3–5, the minor groove width is wider. This is typical of Dickerson-Drew-type DNA duplexes that are packed in the orthorhombic cell. The widening occurs because the duplex accepts the end of a neighboring duplex along the c axis through two hydrogen bonds to form a C:G:G:C quartet. The other end (the residues 8-10) is not widened. However, the corresponding width of every fUG crystal is wider, as discussed later.

3.3. Hydrogen Bonding Schemes of Base Pairs. In every crystal, the two DNA strands are associated to each other through base-pair formations, and all the bases, except for the modified bases and their counter bases positioned on the opposite strands in the duplexes, form the canonical Watson-Crick base pairs. All the pairing geometries of fU residues are shown in Figure 5. As depicted in Figure 1(a), the electron densities clearly show that the fU residues in fUA[2] are paired with the opposite A residues in the Watson-Crick geometry. Those of other fUA crystals also show the same paired structures. These results indicate that the oxidized T residue (fU) still has an ability to form a Watson-Crick-type base pair

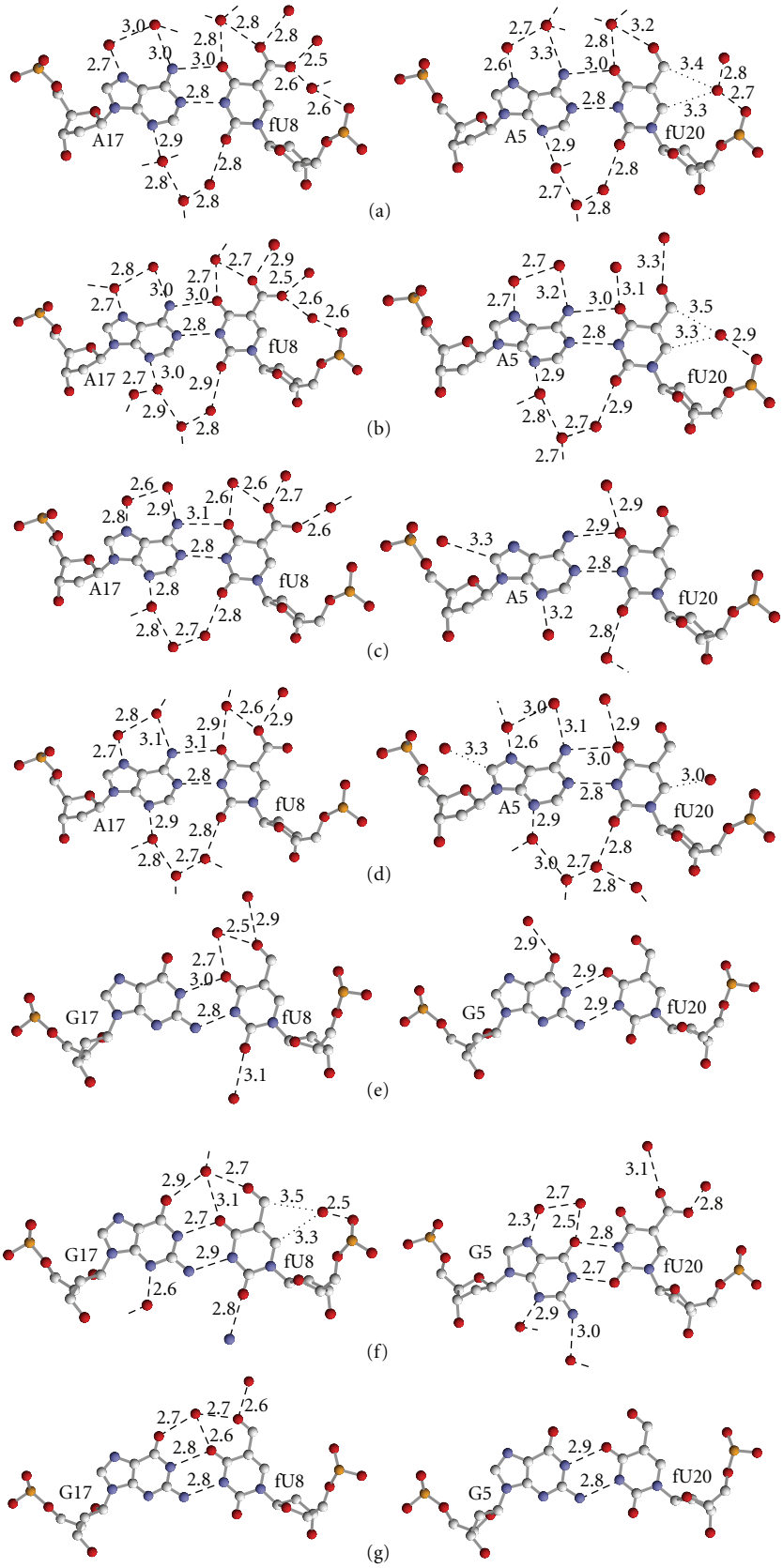


FIGURE 5: The pair formation geometry of fU residues found in (a) fUA[1], (b) fUA[2], (c) fUA[3], (d) fUA[4], (e) fUG[1], (f) fUG[2], and (g) fUG[3]. Broken and dotted lines indicate possible hydrogen bonds and CH$\cdots$O interactions, respectively, and values indicate atomic distances in Å. The pictures are drawn by the program *RASMOL* [28].

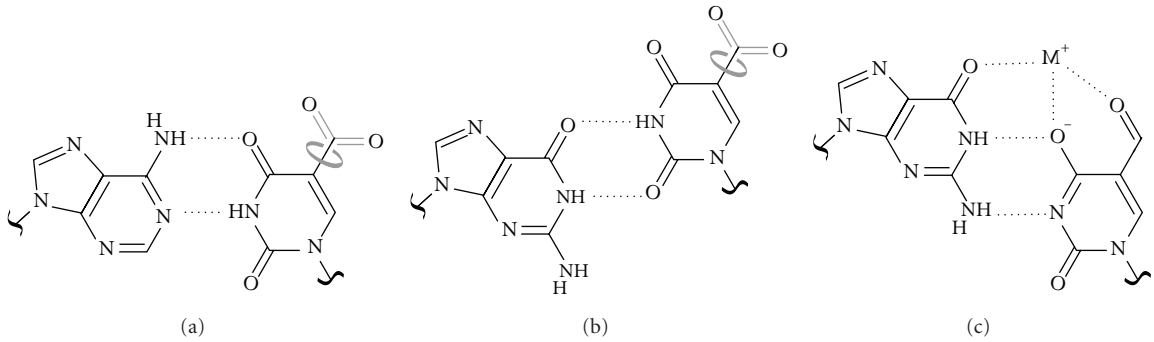


Figure 6: Hydrogen bonding schemes with chemical structures of the observed base pairs: (a) Watson-Crick type, (b) wobble type, and (c) reversed wobble type. M indicates a hydronium or sodium ion.

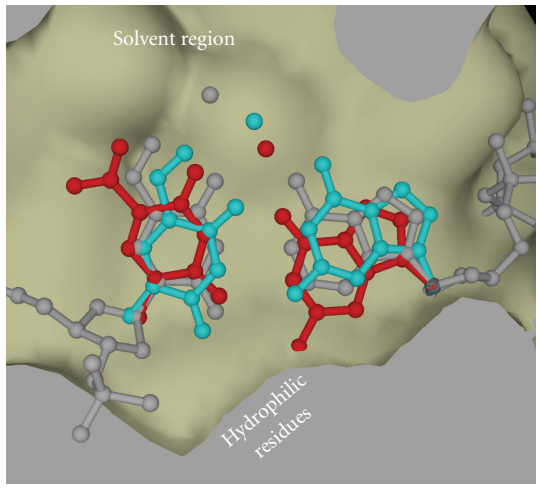


Figure 7: Models of the Watson-Crick-type fU:A (gray) and the reversed wobble fU:G (blue) pairs bound in DNA polymerase I (PDB-ID 3EZ5, see [30]), constructed by fitting the paired bases to the observed pair. The template residues and incoming NTP molecules are bound in the binding pocket of the enzyme, the hydrophilic residues of which may interact to the atoms at the edges of the two bases. The guanine N^2 and C^2 atoms of the wobble fU:G pair (red) are too close to the protein atoms.

with A. The formyl groups of fU could either adopt the *anti* or *syn* conformations depending on the surrounding water structure.

On the other hand, the most interesting features can be seen in the interaction geometries between fU and G in the fUG crystals. In the fUG^2 crystal, the fU20 residue forms a typical wobble pair with G5 through the $N^3H\cdots O^6$ and $O^2\cdots HN^1$ hydrogen bonds, and the formyl group is disordered between *syn* and *anti* conformations. However the fU8 residue forms a new type of pair with the opposite G17 residue through the two hydrogen bonds, $O^4\cdots HN^1$ and $N^3\cdots HN^2$. In the pairing, the G base moves to the major groove side and the fU8 base moves to the minor groove side. Compared to the wobble pair, these bases move in the reverse direction. An atom was located on the electron density map (see Figure 1), and it is within hydrogen distance of the O^4 and O^5 atoms of fU8 and the O^6 atom of G17. As the density of the atoms was of almost the same level as those of other water molecules, a water molecule was temporarily positioned at the peak for further structure refinements. In other words, it seems that a water molecule is trapped in the space surrounded by the three oxygen atoms to stabilize the pair formation. In the fUG^3 crystal, the two fU residues at the different sites also form the respective pairs with the opposite G residues. The geometries are, however, both in the new reversed wobble type, as shown in Figure 1(c). Water molecules are also assigned at the positions similarly surrounded by the three oxygen atoms, as described above, though the corresponding electron density at the fU20 site is rather poor due perhaps to low resolution. Furthermore, in the fUG^1 crystal obtained at pH 8.1 without the help of dyes, the electron density, though again at low resolution, suggests that the two fU residues form pairs with the opposite G residues in a similar manner to those found in fUG^3 obtained with DAPI. Therefore it is considered that the presence of dyes stabilizes duplex structure but does not affect the pairing modes.

4. Discussion

Figure 6 summarizes the pairing modes with their chemical structures, found in the fUA and fUG crystals. It is noteworthy that an oxygen atom always exists at the center almost equidistant from the three surrounding oxygen atoms (O^4 and O^5 of fU and O^6 of G) in the reversed wobble geometry. Because of this pairing, the N^1 atom of fU should be deprotonated and the O^4 atom of fU might be ionized so that the central pocket must accept a hydronium ion instead of a water molecule to stabilize the pair formation. This is consistent with the fact that the fUG^1 crystal was obtained in alkaline state without the help of dyes. The pK_a value decreases from 9.7 to 8.1 in response to the formylation by T oxidization [29]. A sodium ion but not a potassium could also be accommodated, judging from the size of the pocket.

A comparison of the local base-pair parameters at the modified pairs is given in Table 3. As the sequence is basically palindromic, the duplex has a twofold axis at the centre of the duplex perpendicular to the helical axis. Therefore,

the base pairs at fU8 and fU20 exhibit buckle angles and shear distances with signs (+/−) different between the positions related by the twofold symmetry, but propeller twist angles with the same sign though their absolute values are affected by crystal packing. All of the fU:A pairs satisfy these conditions. (The shear values fluctuate around zero in the standard B-form duplex.) The fUG duplexes also adopt this geometrical rule, except for the shear parameter. In the fUG[2] duplex, the two fU:G pairs have the shear values with the same sign (fU8 −2.24 Å and fU20 −2.08 Å). This means that the G17 and G5 bases which are paired with fU8 and fU20 move in the different directions along the X-axis [see the definition in [19]] , that is, G17 shifts up toward the major groove side and G5 shifts down toward the minor groove side (or fU8 moves down toward the minor groove and fU20 moves up toward the major groove). These movements are just ascribed to the reversed wobbling and the normal wobbling, respectively. Another feature of the fU:G reversed wobbling can be seen in the $C^{1'}$-$C^{1'}$ distance which is longer by 0.6 Å, as compared with those of the Watson-Crick pairing. This expansion is also reflected in the wider minor groove widths found in every fU:G duplexes.

The present work has clarified a total of three interaction modes, a Watson-Crick type for fU:A pairing and two wobble types (wobble and reversed wobble) for fU:G pairing. It has been believed that the DNA polymerase accepts base pairs only with the Watson-Crick geometries. The wobble pairing easily occurs between G and T, but the DNA polymerase [30] eliminates such a pair by sieving the shape of the pairs. In order to examine the possibility that a reversed wobble pair could also be accepted by the enzyme the pairing geometry was modeled into the enzyme *in silico*. As seen in Figure 7, the reversed wobble pair is reasonably accommodated without atomic collisions by slightly rotating the pair. On the contrary, a wobble pair, rotated in the opposite direction, would still collide with atoms of the enzyme. Thus, it could be concluded that G forms a pair with fU in the ionized form and A forms a Watson-Crick pair with fU and that the reversed wobble pair is allowed to be bound in the DNA polymerase. This tolerance explains the occurrence of mis-incorporation during replication and leads to pyrimidine transition mutagenesis.

Methoxyamine is also well known as a potent mutagen. N^6-methoxyl-adenine [31, 32] and N^4-methoxyl-cytosine [33–35] residues in the damaged DNAs adopt the imino tautomer to form mismatched base pairs, which mimic the Watson-Crick pair geometries. In contrast, fU derived from thymine oxidation forms a reversed wobble pair, which is also acceptable to the DNA polymerase.

The focus of the present work is on pyrimidine transition mutation and not gene transversion. The atomic mechanism of the latter will be revealed by similar X-ray studies.

Acknowledgments

This work was supported in part by a grant for the RFTF Project from the Japanese Society for the Promotion of Science and by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan. The authors thank N. Igarashi and S. Wakatsuki for facilities and help during data collection.

References

[1] S. S. Wallace, "Enzymatic processing of radiation-induced free radical damage in DNA," *Radiation Research*, vol. 150, no. 5, pp. S60–S79, 1998.

[2] B. Halliwell and O. I. Aruoma, "DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems," *FEBS Letters*, vol. 281, no. 1-2, pp. 9–19, 1991.

[3] M. Yoshida, K. Makino, H. Morita, H. Terato, Y. Ohyama, and H. Ide, "Substrate and mispairing properties of 5-formyl-2′-deoxyuridine 5′-triphosphate assessed by in vitro DNA polymerase reactions," *Nucleic Acids Research*, vol. 25, no. 8, pp. 1570–1577, 1997.

[4] H. Terato, A. Masaoka, M. Kobayashi et al., "Enzymatic repair of 5-formyluracil. II. Mismatch formation between 5-formyluracil and guanine during DNA replication and its recognition by two proteins involved in base excision repair (AlkA) and mismatch repair (MutS)," *The Journal of Biological Chemistry*, vol. 274, no. 35, pp. 25144–25150, 1999.

[5] K. Fujikawa, H. Kamiya, and H. Kasai, "The mutations induced by oxidatively damaged nucleotides, 5-formyl-dUTP and 5-hydroxy-dCTP, in *Escherichia coli*," *Nucleic Acids Research*, vol. 26, no. 20, pp. 4582–4587, 1998.

[6] H. Kamiya, N. Murata-Kamiya, N. Karino, Y. Ueno, A. Matsuda, and H. Kasai, "Induction of T → G and T → A transversions by 5-formyluracil in mammalian cells," *Mutation Research*, vol. 513, no. 1-2, pp. 213–222, 2002.

[7] M. Tsunoda, N. Karino, Y. Ueno, A. Matsuda, and A. Takénaka, "Crystallization and preliminary X-ray analysis of a DNA dodecamer containing 2′-deoxy-5-formyluridine; what is the role of magnesium cation in crystallization of Dickerson-type DNA dodecamers?" *Acta Crystallographica Section D*, vol. 57, no. 2, pp. 345–348, 2001.

[8] M. Tsunoda, J. Kondo, N. Karino, Y. Ueno, A. Matsuda, and A. Takenaka, "Water mediated Dickerson-Drew-type crystal of DNA dodecamer containing 2′-deoxy-5-formyluridine," *Biophysical Chemistry*, vol. 95, no. 3, pp. 227–233, 2002.

[9] H. Sugiyama, S. Matsuda, K. Kino, Q.-M. Zhang, S. Yonei, and I. Saito, "New synthetic method of 5-formyluracil-containing oligonucleotides and their melting behavior," *Tetrahedron Letters*, vol. 37, no. 50, pp. 9067–9070, 1996.

[10] Z. Otwinowski and W. Minor, "Processing of X-ray diffraction data collected in oscillation mode," *Methods in Enzymology*, vol. 276, pp. 307–326, 1997.

[11] M. G. Rossmann and C. G. van Beek, "Data processing," *Acta Crystallographica Section D*, vol. 55, no. 10, pp. 1631–1640, 1999.

[12] J. Navaza and P. Saludjian, "AMoRe: an automated package for molecular replacement," *Acta Crystallographica A*, vol. 50, pp. 157–163, 1994.

[13] X. Shui, L. McFail-Isom, G. G. Hu, and L. D. Williams, "The B-DNA dodecamer at high resolution reveals a spine of water on sodium," *Biochemistry*, vol. 37, no. 23, pp. 8341–8355, 1998.

[14] P. Emsley and K. Cowtan, "Coot: model-building tools for molecular graphics," *Acta Crystallographica Section D*, vol. 60, no. 12, pp. 2126–2132, 2004.

[15] "Collaborative Computational Project," Number 4, 1994.

[16] G. N. Murshudov, A. A. Vagin, and E. J. Dodson, "Refinement of macromolecular structures by the maximum-likelihood

method," *Acta Crystallographica Section D*, vol. 53, no. 3, pp. 240–255, 1997.

[17] A. T. Brünger, P. D. Adams, G. M. Clore et al., "Crystallography & NMR system: a new software suite for macromolecular structure determination," *Acta Crystallographica Section D*, vol. 54, no. 5, pp. 905–921, 1998.

[18] A. Philippsen, "DINO: Visualizing Structural Biology," 2002, http://www.dino3d.org/.

[19] W. K. Olson, M. Bansal, S. K. Burley et al., "A standard reference frame for the description of nucleic acid base-pair geometry," *Journal of Molecular Biology*, vol. 313, no. 1, pp. 229–237, 2001.

[20] X.-J. Lu and W. K. Olson, "3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures," *Nucleic Acids Research*, vol. 31, no. 17, pp. 5108–5121, 2003.

[21] E. C. M. Juan, S. Shimizu, X. Ma et al., "Insights into the DNA stabilizing contributions of a bicyclic cytosine analogue: crystal structures of DNA duplexes containing 7,8-dihydropyrido[2,3-*d*]pyrimidin-2-one," *Nucleic Acids Research*, vol. 2010, pp. 1–9, 2010.

[22] M. K. Teng, N. Usman, C. A. Frederick, and A. H. Wang, "The molecular structure of the complex of Hoechst 33258 and the DNA dodecamer d(CGCGAATTCGCG)," *Nucleic Acids Research*, vol. 16, no. 6, pp. 2671–2690, 1988.

[23] J. R. Quintana, A. A. Lipanov, and R. E. Dickerson, "Low-temperature crystallographic analyses of the binding of Hoechst 33258 to the double-helical DNA dodecamer C-G-C-G-A-A-T-T-C-G-C-G," *Biochemistry*, vol. 30, no. 42, pp. 10294–10306, 1991.

[24] M. Sriram, G. A. van der Marel, H. L. P. F. Roelen, J. H. Van Boom, and A. H.-J. Wang, "Conformation of B-DNA containing O6-ethyl-G-C base pairs stabilized by minor groove binding drugs: molecular structure of d(CGC[e6G]AATTCGCG complexed with Hoechst 33258 or Hoechst 33342," *EMBO Journal*, vol. 11, no. 1, pp. 225–232, 1992.

[25] G. R. Clark, C. J. Squire, E. J. Gray, W. Leupin, and S. Neidle, "Designer DNA-binding drugs: the crystal structure of a meta-hydroxy analogue of Hoechst 33258 bound to $d(CGCGAATTCGCG)_2$," *Nucleic Acids Research*, vol. 24, no. 24, pp. 4882–4889, 1996.

[26] T. A. Larsen, D. S. Goodsell, D. Cascio, K. Grzeskowiak, and R. E. Dickerson, "The structure of DAPI bound to DNA," *Journal of Biomolecular Structure and Dynamics*, vol. 7, no. 3, pp. 477–491, 1989.

[27] D. Vlieghe, J. Sponer, and L. Van Meervelt, "Crystal structure of d(GGCCAATTGG) complexed with DAPI reveals novel binding mode," *Biochemistry*, vol. 38, no. 50, pp. 16443–16451, 1999.

[28] R. A. Sayle and E. J. Milner-White, "RASMOL: biomolecular graphics for all," *Trends in Biochemical Sciences*, vol. 20, no. 9, pp. 374–376, 1995.

[29] E. J. Privat and L. C. Sowers, "A proposed mechanism for the mutagenicity of 5-formyluracil," *Mutation Research*, vol. 354, no. 2, pp. 151–156, 1996.

[30] A. A. Golosov, J. J. Warren, L. S. Beese, and M. Karplus, "The mechanism of the translocation step in DNA replication by DNA polymerase I: a computer simulation analysis," vol. 18, no. 1, pp. 83–93, 2010.

[31] T. Chatake, A. Ono, Y. Ueno, A. Matsuda, and A. Takénaka, "Crystallographic studies on damaged DNAs. I. An N^6-methoxyadenine residue forms a Watson-Crick pair with a cytosine residue in a B-DNA duplex," *Journal of Molecular Biology*, vol. 294, no. 5, pp. 1215–1222, 1999.

[32] T. Chatake, T. Hikima, A. Ono, Y. Ueno, A. Matsuda, and A. Takénaka, "Crystallographic studies on damaged DNAs. II. N^6-methoxyadenine can present two alternate faces for Watson-Crick base-pairing, leading to pyrimidine transition mutagenesis," *Journal of Molecular Biology*, vol. 294, no. 5, pp. 1223–1230, 1999.

[33] M. T. Hossain, T. Sunami, M. Tsunoda et al., "Crystallographic studies on damaged DNAsIV. N^4-methoxycytosine shows a second face for Watson-Crick base-pairing, leading to purine transition mutagenesis," *Nucleic Acids Research*, vol. 29, no. 19, pp. 3949–3954, 2001.

[34] M. T. Hossain, T. Chatake, T. Hikima et al., "Crystallographic studies on damaged DNAs: III. N^4-methoxycytosine can form both Watson-Crick type and wobbled base pairs in a B-form duplex," *Journal of Biochemistry*, vol. 130, no. 1, pp. 9–12, 2001.

[35] M. Tofazzal Hossain, J. Kondo, Y. Ueno, A. Matsuda, and A. Takénaka, "X-Ray analysis of d(CGCGAATTXGCG)2 containing a 2′-deoxy-N^4-methoxycytosine residue at X: a characteristic pattern of sugar puckers in the crystalline state of the Dickerson-Drew type DNA dodecamers," *Biophysical Chemistry*, vol. 95, no. 1, pp. 69–77, 2002.