Crystal Structure of a DNA Duplex Containing 8-Hydroxydeoxyguanine–Adenine Base Pairs[†,‡]

Katherine E. McAuley-Hecht,[§] Gordon A. Leonard,[∥] Neil J. Gibson,[§] James B. Thomson,[§] William P. Watson,[⊥] William N. Hunter,[∥] and Tom Brown*[,§]

Edinburgh Centre for Molecular Recognition, Department of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JJ, U.K., Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, U.K., and Shell Research Ltd., Sittingbourne Research Centre, Sittingbourne, Kent ME9 8AG, U.K.

Received February 23, 1994; Revised Manuscript Received May 26, 1994[⊗]

ABSTRACT: The crystal structure of the oligonucleotide d(CGC*A*AATT*O8G*GCG), containing the chemically modified base 8-hydroxydeoxyguanine (O8G), has been determined at 2.5-Å resolution and refined to a crystallographic *R*-factor of 16.8%. The B-type DNA helix contains standard Watson–Crick base pairs except at the mismatch sites, where O8G adopts a *syn* conformation and forms hydrogen bonds to adenine in the *anti* conformation. The thermodynamic stability of the duplex was found by UV melting techniques to be intermediate between the native oligonucleotide d(CGC*A*AATT*T*GCG) and an oligonucleotide containing A·G mispairs d(CGC*A*AATT*G*GCG). Comparison of the structure of the O8G(*syn*)·A(*anti*) base pair with those of Watson–Crick base pairs has given a reason why O8G·A base pairs are not well repaired by DNA proofreading enzymes.

8-Hydroxydeoxyguanosine (or 8-oxo-7,8-dihydro-2′-deoxyguanosine) (O8G) is produced *in vivo* from the hydroxylation of deoxyguanosine residues by oxygen radicals and is the most common product of oxidative damage to DNA (Kasai & Nishimura, 1984; Shigenaga et al., 1989). This modified base can exist in several tautomeric forms but, under physiological conditions, the 6,8-diketoguanine form predominates (Aida & Nishimura, 1987; Culp et al., 1989).

The study of O8G is of some importance since its presence in DNA has been linked to several diseases, e.g., rheumatoid arthritis (RA), systemic lupus erythomatosus (SLE), and cancer. The association of O8G and the autoimmune disease RA has been demonstrated by the high levels (five times greater than that of controls) of O8G found in the urine of patients with this disease. Patients with SLE do not have increased urine O8G levels but have high levels of antibodies to oxidatively modified DNA (Herbert et al., 1993). There is a direct correlation between conditions that lead to cancer, such as ionizing irradiation or exposure to chemical carcinogens and high levels of O8G in DNA (Floyd, 1990).

O8G is promutagenic, and during replication of a DNA template containing an O8G residue, the polymerase may insert dA instead of dC opposite the mutated base, with the ratio of dA/dC incorporation being dependent on the DNA polymerase involved (Shibutani et al., 1991). Neither O8G·C nor O8G·A base pairs are corrected by the polymerase editing function, and so O8G is mutagenic and potentially carcinogenic since dG to dT transversion mutations will result from the misincorporation of dA into the new DNA strand.

Because this primary lesion occurs with high frequency, several dedicated repair enzymes have evolved specifically to correct O8G·A base pairs. In *Escherichia coli*, the MutM protein removes O8G lesions from DNA, while the MutY protein excises the adenine base from the O8G·A mispair (Michaels et al., 1992). As a further defense, the MutT protein eliminates 8-hydroxy-dGTP from the nucleotide pool to prevent O8G being incorporated opposite adenine during replication (Maki & Sekiguchi, 1992). The 8-hydroxyguanine repair enzymes found in human cells are effective at removing O8G from O8G·C, O8G·T, and O8G·G base pairs but not at correcting O8G·A base pairs (Bessho et al., 1993).

Structural studies of DNA duplexes containing O8G provide insight into why DNA polymerases are unable to recognize and correct these lesions. Thermodynamic data are also important to the understanding of DNA repair mechanisms. Here we describe the crystal structure and stability of a dodecamer with two A(*anti*)·O8G(*syn*) base pairs.

EXPERIMENTAL PROCEDURES

Oligonucleotide Synthesis. The synthesis of the 8-hydroxyguanine phosphoramidite is illustrated in Figure 1. 7,8-Dihydro-8-oxo-N^2-isobutyryl-2′-deoxyguanosine (**4**) was prepared from 2′-deoxyguanosine (**1**) and incorporated into DNA by modifications of the methods by Bodepudi et al. (1992) and Roelen et al. (1991). The phosphoramidite **6** could be prepared by treatment of the alcohol **5** with 2-cyanoethyl *N,N,N′,N′*-tetraisopropylphosphorodiamidite (**7**) in the presence of diisopropylammonium tetrazolide followed by flash column chromatography. Protection of the lactam functionality of the alcohol **5** was found to be unnecessary. The phosphoramidite **6** coupled to the growing DNA chain with an efficiency of >96%. All syntheses were carried out "trityl on" on an ABI 394A automated DNA synthesizer. The oligonucleotide was cleaved from the resin and deprotected for 20 h at 65 °C in a mixture of ethanethiol (0.25 M) in concentrated aqueous ammonia (30% v/v). Purification was by trityl-on HPLC followed by cleavage of the trityl group (80% aqueous acetic acid for 1 h) and further reverse-phase HPLC purifications of the detritylated product. The fully purified oligonucleotide was then lyophilized and stored at 4 °C before use in crystallizations.

[†] Supported by The Wellcome Trust, the Science and Engineering Research Council, and Shell Research Ltd.

[‡] Final refined coordinates and structure factors have been deposited with the Brookhaven Protein Data Bank (file name 178D).

* Correspondence should be addressed to this author.
[§] Department of Chemistry, University of Edinburgh.
[∥] Department of Chemistry, University of Manchester.
[⊥] Shell Research Ltd.

[⊗] Abstract published in *Advance ACS Abstracts,* July 15, 1994.

0006-2960/94/0433-10266$04.50/0

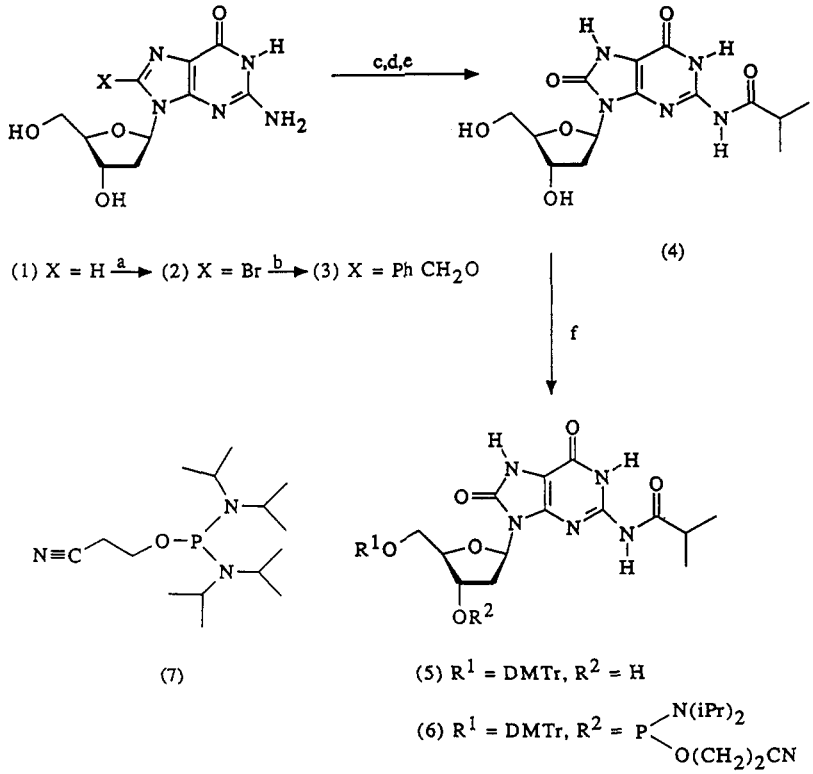

FIGURE 1: Scheme illustrating the steps involved in the synthesis of the 8-hydroxydeoxyguanine phosphoramidite: a, Br_2, H_2O; b, $PhCH_2OH/PhCH_2O^-$, dimethyl sulfoxide, 60 °C; c, $Me_2CHCOCl$, pyridine; d, pyridine, methanol, 2 M KOH; e, H_2 (60 atm), Pd/C; f, dimethoxytrityl chloride, pyridine; g, 7, diisopropylammonium tetrazolide, acetonitrile.

Crystallization. Needlelike crystals were grown at 4 °C, from 20-μL sitting droplets, equilibrated against a reservoir of 100% 2-methyl-2,4-pentanediol. The crystallization solution contained approximately 1 mM oligonucleotide, 20 mM magnesium chloride, 8% (v/v) 2-methyl-2,4-pentanediol, and 1 mM spermine tetrahydrochloride in 12 mM sodium cacodylate buffer (pH 6.5).

Data Collection. A crystal of dimensions 0.8 × 0.1 × 0.1 mm, mounted in a glass capillary, was used to collect data to 2.5-Å resolution on a Rigaku RU200 rotating anode generator operating at 50 kV, 140 mA, with a crystal to detector distance of 138 mm. Orthorhombic unit cell dimensions of a = 24.67 Å, b = 40.49 and c = 65.14 Å were determined from three still photographs taken at 45° intervals with an exposure time of 20 min. A total of 9923 reflections were measured from 30 oscillation frames at 3° intervals and 75-min exposure times. Of these, 5480 with $I \geq \sigma(I)$ were accepted. Autoindexing, data scaling, the reduction were carried out using the software PROCESS (Higashi, 1990). The 5480 accepted data were reduced to 1965 unique reflections (redundancy factor: 2.8) representing 81% of the total possible reflections to 2.5-Å resolution. The merging R value was 6.8%.

Structure Refinement. A rigid body refinement was carried out using a modified version of SHELX (Sheldrick, 1976), with the coordinates of d(CGCGAATTCGCG) as a starting model (Drew et al., 1980). The refinement proceeded in steps of 1 Å, starting with data in the region 10–6 Å, and gradually increasing the resolution to 3 Å, at which point the R-factor was 32% for 1263 reflections.

A restrained least-squares refinement followed, using the program NUCLSQ (Westhof et al., 1985). The nucleotides at positions 4, 9, 16, and 21 were removed from the structure factor calculations, and the refinement was continued with the 2σ data (1965 reflections) in the resolution range 7–2.5 Å to give an R-factor of 38%. $2F_o - F_c$ and $F_o - F_c$ density maps were calculated and displayed on an Evans and Sutherland graphics system using FRODO (Jones, 1978). The density maps clearly indicated that the bases at positions 9 and 21 should be in a *syn* conformation, while the bases at positions 4 and 16 should be in an *anti* conformation. The DNA model was modified to include two A(*anti*)·O8G(*syn*) base pairs, and these were then included in the refinement, which converged at R = 0.25. Addition of 72 water molecules, located from the difference density maps coupled with the restrained refinement of individual isotropic thermal parameters, led to a final R-factor of 16.8% for 1843 2$\sigma(F)$ reflections to a resolution of 2.5 Å. The geometry of the final model is excellent with root mean square (rms) deviations from ideality of 0.006 Å for bond distances in the sugar/base groups and 0.005 Å for bond distances in the phosphate groups. For angle distances the rms deviations are 0.014 and 0.010 Å, respectively. An example of the fit of the final model to the electron density is given in Figure 2. Final refined coordinates and structure factors have been deposited with the Brookhaven Protein Data Bank.

UV Melting Studies. UV melting curves were measured with a Perkin-Elmer Lambda 15 UV spectrophotometer equipped with a Peltier heating block and controlled by an IBM PS2 computer. Samples of the DNA dodecamers were dissolved in a buffer consisting of 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 1 mM EDTA, at concentrations of 10 μM. Heating was performed over a temperature range of 5–70 °C at a rate of 0.9 deg/min, and the data were processed with the PECSS software package.

RESULTS

Conformational Features. The nucleotides are labeled from C1 to G12 in the 5′–3′ direction on strand 1 and from C13 to G24 in the 5′–3′ direction on strand 2. The overall structure of d(CGC***A***AATT***O8G***GCG) is typical of a standard B-DNA-type helix. There are 9.9 base pairs per helix turn, the average rise between base pairs is 3.3 Å, the average distance

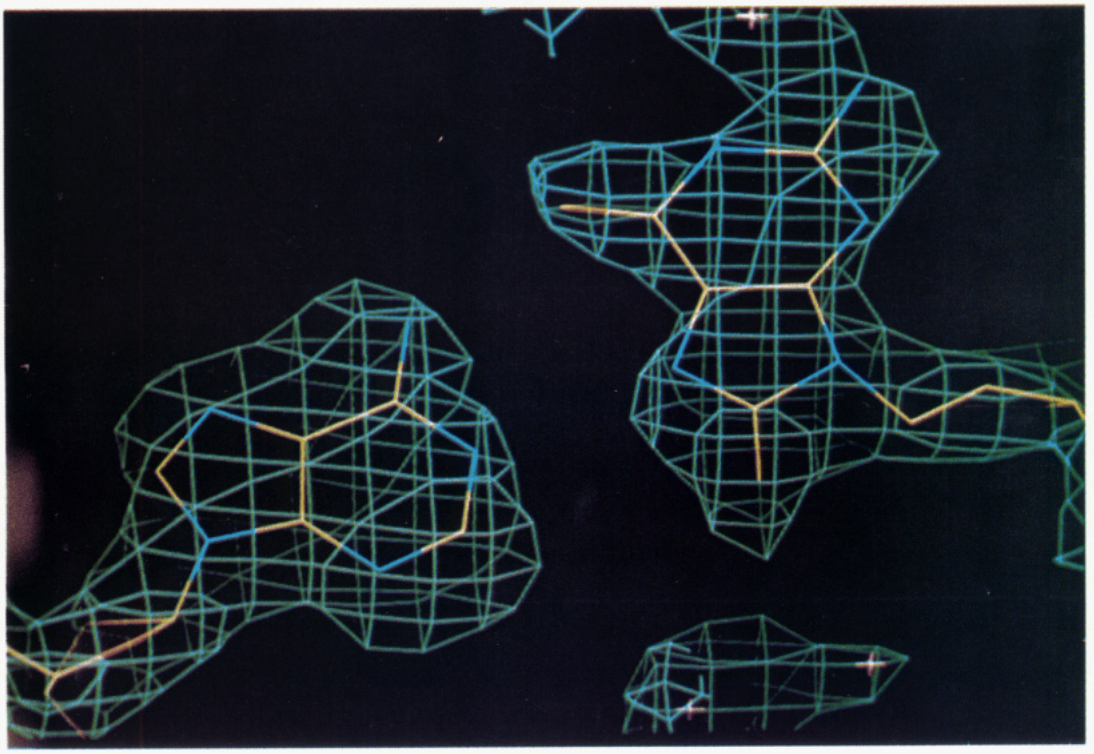

FIGURE 2: O8G9·A16 base pair superimposed on a $2F_o - F_c$ electron density map (in blue). The red crosses represent water molecules.

Table 1: Distances and Angles for the Proposed Hydrogen Bonds between O8G and Adenine at Base Pairs 9·16 and 4·21

measurement	base pair 9·16	base pair 4·21
distances (Å)		
O6(O8G)–N6(A)	2.67	3.06
N7(O8G)–N1(A)	2.48	2.70
angles (deg)		
C6(O8G)–O6(O8G)–N6(A)	131.0	116.4
O6(O8G)–N6(A)–C6(A)	123.3	140.8
C5(O8G)–N7(O8G)–N1(A)	128.2	147.6
N7(O8G)–N1(A)–C6(A)	129.4	108.1

between adjacent phosphates is 6.7 Å, and the sugar puckers are mainly C2′-endo. Therefore, the gross structure is not significantly different from that of the native oligonucleotide d(CGC*G*AATT*C*GCG) (Drew et al., 1980). The A(*anti*)·O8G(*syn*) base pairs are accommodated into the duplex without noticeable distortion to the sugar–phosphate backbone. The mispairs are slightly asymmetrical with angles λ_1 and λ_2 (Hunter et al., 1986a) of 50.2° and 46.3° for A4·O8G21 and of 41.9° and 55.5° for O8G9·A16, respectively. The C1′–C1′ distances of 10.54 and 10.40 Å are standard for B-DNA and compare to an average C1′–C1′ distance of 10.42 Å for the remaining base pairs.

Base-Pair Geometry. The structure of the O8G9·A16 base pair is illustrated in Figures 2 and 3 (the A4·O8G21 base pair is almost identical). Hoogsten-like base pairs are formed, with the 8-hydroxydeoxyguanine bases in the *syn* conformation and A4 and A16 in the *anti* conformation. For both base pairs the angles and distances between C(6)–N(6)–adenine and O(6)–C(6)–O8G and C(6)–N(1)–adenine to N(7)–C(5)–O8G indicate hydrogen bond formation (Table 1). The oxygen atoms at position 8 of the O8G base do not appear, at first sight, to be involved in hydrogen-bonding interactions. However, a weak C–H···O hydrogen bond may exist to adenine H(2).

Thermal Stability. The melting temperatures (T_m) determined in the UV melting studies are summarized in Table 2. They indicate that the oligonucleotide containing two 8-hydroxydeoxyguanine–adenine base pairs is destabilized by only 6 °C compared to the native sequence d(CGC*A*AATT*T*GCG). This is a minor destabilization relative to other mismatch sequences. For example, a much larger destabilization is observed for the dodecamer d(CGC*A*AATT*G*GCG), for which the T_m is depressed by more than 35 °C at this pH (Table 2). The oligonucleotide containing the C·O8G base pairs melts at a lower temperature than d(CGC*A*AATT*O8G*GCG) and is significantly destablilized compared to the oligonucleotide d(CGC*C*AATT*G*GCG), probably because of steric effects (see below).

Table 2: UV Melting Temperatures for Dodecamers of the Type d(CGCXAATTYGCG) at Concentrations of 10 μM in 10 mM sodium phosphate (pH 7.0), 100 mM NaCl, and 1 mM EDTA

oligonucleotide	melting temp (K)
CGC*A*AATT*T*GCG	326.9
CGC*A*AATT*O8G*GCG	321.4
CGC*C*AATT*O8G*GCG	313.5
CGC*C*AATT*G*GCG	330.4
CGC*A*AATT*G*GCG	292.3

DISCUSSION

O8G forms base pairs with both dC and dA, depending on whether the base adopts an *anti* or a *syn* conformation about the glycosidic bond. The *syn* conformation is favored by monomeric O8G owing to steric hindrance between O8 and the 4′-oxygen atom of the deoxyribose moiety (Uesugi & Ikehara, 1977). However, O8G(*anti*)·C(*anti*) base pairs can form as shown by the NMR studies of d(CGC*O8G*AATT*C*GCG) (Oda et al., 1991). The structure of an A(*anti*)·O8G(*syn*) base pair in solution has also been determined by NMR (Kouchadkjian et al., 1991), and the major species in the spectrum is an A(*anti*)·O8G(*syn*) base pair which is identical to the mispair in the present crystal structure. The structure of the A(*anti*)·O8G(*syn*) base pair explains why dA can be incorporated opposite O8G during replication: O8G in the *syn* conformation exhibits the required pattern of hydrogen-bonding acceptor/donor groups and the

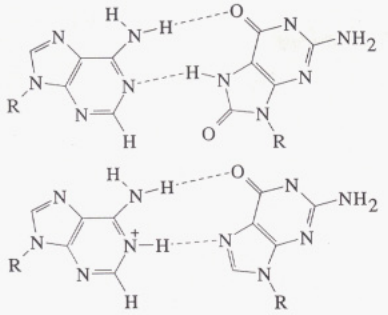

FIGURE 3: (Top) Representation of the structure of the A(*anti*)·O8G-(*syn*) base pair. The dashed lines represent hydrogen bonds. (Bottom) Same representation for the unmodified A(*anti*)·G(*syn*) base pair.

correct geometry to form hydrogen bonds to, and therefore a stable base pair with, adenine.

Purine–purine mispairs can take several conformational forms; both *anti–anti* and *anti–syn* conformations have been identified by X-ray crystallography and NMR [e.g., Gao et al. (1988), Lane et al. (1991), Brown et al. (1986), and Kan et al. (1983)]. The conformation of the G·A mismatch is sequence dependent, with G(*anti*)·A(*anti*) base pairs in the sequence CCAA*GA*TTGG (Privé et al., 1987), G(*anti*)·A(*syn*) base pairs in the crystal structure of CGC*G*AATT*A*GCG (Hunter et al., 1986b), and G(*syn*)·A-(*anti*) base pairs in CGC*A*AATT*G*GCG (Leonard et al., 1990). This latter duplex contains A(*anti*)·G(*syn*) base pairs that are analogous to the A(*anti*)·O8G(*syn*) base pairs. The main difference between these two duplexes lies in their thermal stabilities: at physiological pH the A·G duplex has a melting temperature which is almost 30° lower than the melting temperature of the A(*anti*)·O8G(*syn*) duplex (Table 2). The reason for this difference is that the adenine base must be protonated at N1 to form a second hydrogen bond to the unmodified guanine base (Figure 3). The duplex is therefore unstable at higher pH values but becomes relatively stable at low pH (Brown et al., 1990). In contrast, O8G exists as the 6,8-diketo tautomer with a proton donor at N7, and so protonation is not required for this base to form two hydrogen bonds to adenine.

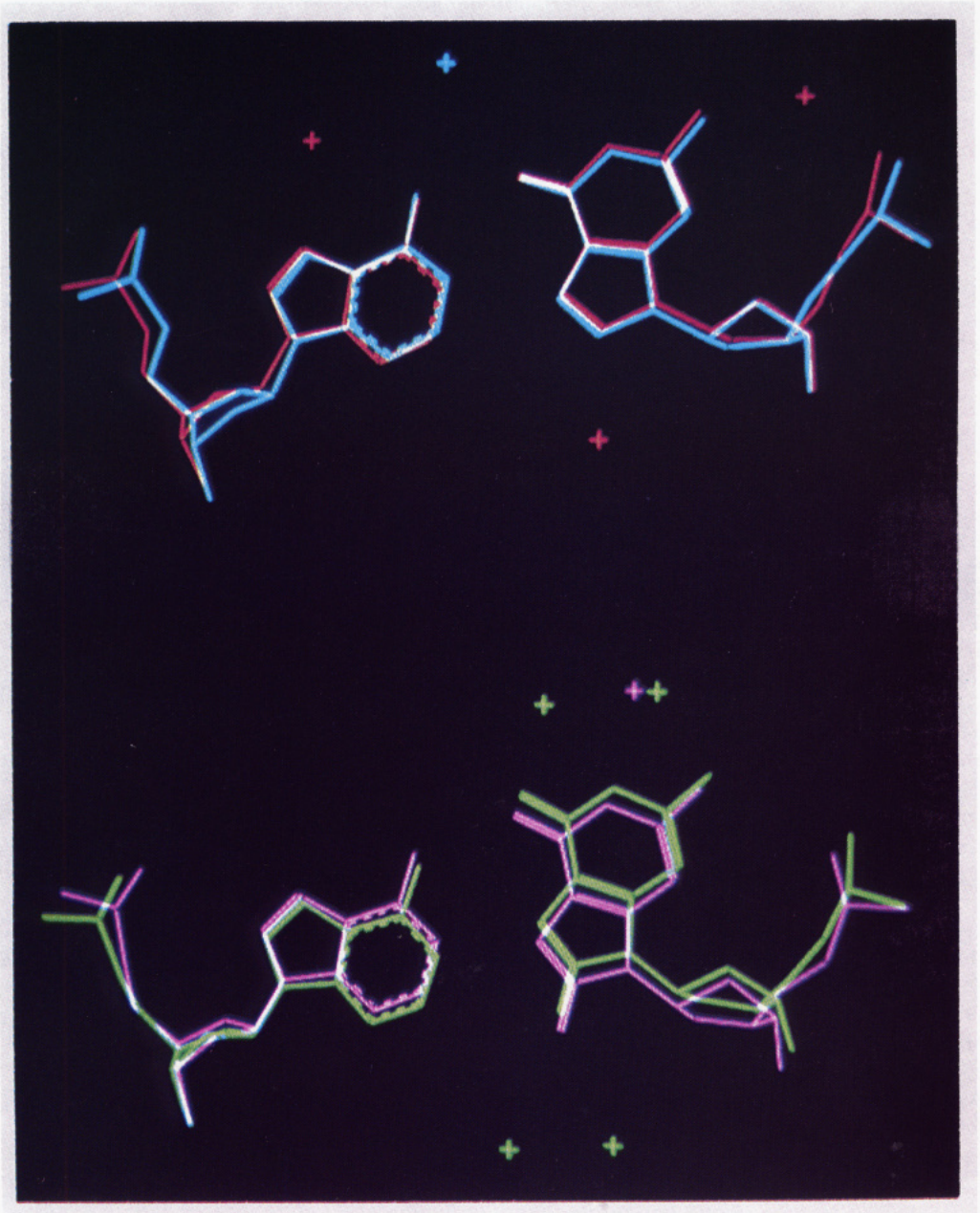

FIGURE 4: (Top) Superposition of A4·G21 (red) and A16·G9 (cyan) from the structure d(CGC*A*AATT*G*GCG) (Leonard et al., 1990), with neighboring water molecules. (Bottom) Superposition of A4·O8G21 (magenta) and A16·O8G9 (green) from the structure d(CGC*A*AATT*O8G*GCG), with neighboring water molecules.

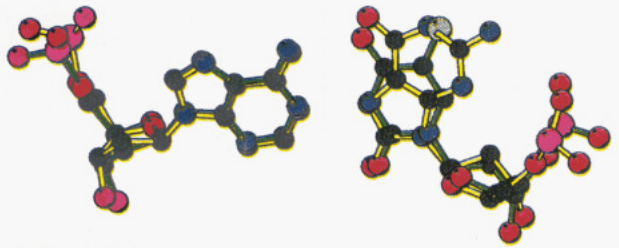

FIGURE 5: Superposition of the A(*anti*)·O8G(*syn*) base pair (yellow bonds) on the A5·T20 base pair (green bonds) from the structure of d(CGCGAATT(ɛdA)CGC) (Leonard et al., 1994). The superposition involved the functional groups related to hydrogen bonding, the nitrogen atoms in the glycosidic bonds, and the N3, N7 atoms of the adenine bases. Carbon atoms are colored black, oxygen red, nitrogen blue, and phosphorus pink. The methyl group on the thymidine base is in gray.

The stability of the mismatch may also be affected by base-pair hydration, and Figure 4 illustrates the hydration pattern of the A(*anti*)·G(*syn*) and the A(*anti*)·O8G(*syn*) base pairs. The positioning of water molecules around the mismatches is different in each case.

DNA replication has a very low error rate. In addition to its polymerase action, DNA polymerase I also has a 3′–5′ exonuclease function which excises mismatched nucleotides from the growing DNA strand. Although the structures of the A(*anti*)·G(*syn*) mispair and the A(*anti*)·O8G(*syn*) mispair are very similar, the former is corrected by DNA polymerase (although with low efficiency) while the latter is not. This difference in the ability of the exonuclease to repair these errors might be explained by the different thermal stabilities of the mispairs since the lower melting temperatures of mismatched bases are important in promoting the 3′–5′ exonuclease activity of DNA polymerase. Therefore, the A(*syn*)·O8G(*anti*) base pair may not be recognized as an error because of its relatively high thermodynamic stability.

Other factors which may be significant in error detection during replication are the asymmetry of the base pair and the disposition of H-bonding donors and acceptors in the minor and major grooves of the DNA duplex. Structural analysis of various mispairs suggests that the more asymmetrical base pairs are corrected with higher efficiency than mispairs that resemble symmetric Watson–Crick pairs more closely (Hunter et al., 1986a). The A(*anti*)·O8G(*syn*) base pair is less asymmetric about its glycosidic bonds than most mispairs, and this may be another reason why it is not removed by DNA repair mechanisms.

The A(*anti*)·O8G(*syn*) base pair is structurally similar to an A·T base pair, even to the positioning of a carbonyl oxygen in the minor groove of the DNA. This means that the A(*anti*)·O8G(*syn*) base pair may not be distinguished from an A·T base pair in the minor groove: only the major groove side can act as a recognition site for repair enzymes. This may not be a major factor since most sequence-specific recognition is thought to occur in the major groove, the minor groove being too narrow to accommodate an α-helix and inaccessible to protein side chains. There is still the possibility that recognition to the minor groove could occur via water-mediated interactions. A superposition of an A·T base pair with an A(*anti*)·O8G(*syn*) base pair is given in Figure 5. This illustrates the close structural similarity between the A·T and the A(*anti*)·O8G(*syn*) base pairs, particularly in the minor groove but also in the backbone regions.

The overall geometry of the A(*anti*)·O8G(*syn*) base pair, its high thermodynamic stability, and the presence of a carbonyl oxygen in the minor groove side result in a base pair that is very difficult to distinguish from an A·T base pair. This similarity accounts for the failure of DNA polymerase I to repair this lesion and that specialized repair enzymes are required to remove O8G bases in DNA.

REFERENCES

Aida, M., & Nishimura, S. (1987) *Mutat. Res. 192*, 83–89.

Bessho, T., Tano, K., Kasai, H., Ohtsuka, E., & Nishimura, S. (1993) *J. Biol. Chem. 268*, 19416–19421.

Bodepudi, V., Shibutani, S., & Johnson, F. (1992) *Chem. Res. Toxicol. 5*, 608–617.

Brown, T., Hunter, W. N., Kneale, G., & Kennard, O. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2402–2406.

Brown, T., Leonard, G. A., Booth, E. D., & Kneale, G. (1990) *J. Mol. Biol. 212*, 437–440.

Culp, S. J., Cho, B. P., Kadlubar, F. F., & Evans, F. E. (1989) *Chem. Res. Toxicol. 2*, 416–422.

Drew, H. R., Wing, R. M., Takano, T., Broka, C., Takana, S., Itakura, K., & Dickerson, R. E. (1980) *Nature 287*, 755–758.

Floyd, R. A. (1990) *Carcinogenesis 11*, 1447–1450.

Gao, X., & Patel, D. J. (1988) *J. Am. Chem. Soc. 110*, 5178–5182.

Herbert, K. E., Griffiths, H. R., Mistry, N., Nichol, F. E., Samanta, A. K., Feehally, J., Emery, P., & Lunec, J. (1993) *Br. J. Rheumatol. 32* (Suppl. 2), 29.

Higashi, T. (1990) *PROCESS: A program for indexing and processing R-AXIS imaging plate data*, Rigaku Corp., Tokyo, Japan.

Hunter, W. N., Brown, T., Anand, N. N., & Kennard, O. (1986a) *Nature 320*, 552–555.

Hunter, W. N., Brown, T., & Kennard, O. (1986b) *J. Biomol. Struct. Dyn. 4*, 173–191.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kan, L.-S., Chandrasegaran, S., Ulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263–4265.

Kasai, H., & Nishimura, S. (1984) *Nucleic Acids Res. 12*, 2137–2145.

Kouchakdjian, M., Bodepudi, V., Shibutani, S., Eisenberg, M., Johnson, F., Grollman, A. P., & Patel, D. J. (1991) *Biochemistry 30*, 1403–1412.

Leonard, G. A., Booth, E. D., & Brown, T. (1990) *Nucleic Acids Res. 18*, 5617–5623.

Leonard, G. A., McAuley-Hecht, K. E., Gibson, N. J., Brown, T., Watson, W. P., & Hunter, W. N. (1994) *Biochemistry 33*, 4755–4761.

Maki, H., & Sekiguchi, M. (1992) *Nature 355*, 273–275.

Michaels, M. L., Cruz, C., Grollman, A. P., & Miller, J. H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7022–7025.

Oda, Y., Uesugi, S., Ikehara, M., Nishimura, S., Kawase, Y., Ishikawa, H., Inoue, H., & Ohtsuka, E. (1991) *Nucleic Acids Res. 19*, 1407–1412.

Privé, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M., & Dickerson, R. E. (1987) *Science 238*, 498–504.

Roelen, H. C. P. F., Saris, C. P., Brugghe, H. F., van den Elst, H., Westra, J. G., van der Marel, G. A., & van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 4361–4369.

Sheldrick, G. M. (1976) *SHELX76 System of Computer Programs*, University of Cambridge, Cambridge, U.K.

Shibutani, S., Takeshita, M., & Grollman, A. P. (1991) *Nature 349*, 431–434.

Shigenaga, M. K., Gimenco, C. J., & Ames, B. N. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9697–9701.

Skelly, J. V., Edwards, K. J., Jenkins, T. C., & Neidle, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 804–808.

Uesugi, S., & Ikehara, M. (1977) *J. Am. Chem. Soc. 99*, 3250–3253.

Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., & Swann, P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6693–6697.

Westhof, E., Dumas, P., & Moras, D. J. (1985) *J. Mol. Biol. 184*, 119–145.