Volume 14 Number 4 1986 Nucleic Acids Research

Structural characterisation of the bromouracil.guanine base pair mismatch in a Z-DNA fragment

T.Brown, G.Kneale, W.N.Hunter and O.Kennard

University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK

Received 17 December 1985; Revised and Accepted 22 January 1986

ABSTRACT

The deoxyoligonucleotide d(^{Br}U-G-C-G-C-G) was crystallised at pH 8.2 and its structure analysed by X-ray diffraction. The unit cell, of dimensions a=17.94, b=30.85, c=49.94Å contains four DNA duplexes in space group $P2_12_12_1$. The duplexes are in the Z conformation, with four Watson-Crick G.C base pairs and two ^{Br}U.G base pairs. The structure was refined to an R factor of 0.16 at a resolution of 2.2Å with 64 solvent molecules located. The ^{Br}U.G base pair mismatch is of the wobble type, with both bases in the major tautomer form and hydrogen bonds linking O-2 of ^{Br}U with N-1 of G and N3 of ^{Br}U with O-6 of G. There is no indication of the presence of ionised base pairs, in spite of the high pH of crystallisation. The results are discussed in terms of the mutagenic properties of 5- bromouracil.

INTRODUCTION

The mutagenic base, 5-bromouracil, is an analogue of thymine and as such is incorporated at about the same rate by DNA polymerase opposite template adenine. Bromouracil, however, has a much greater tendency than thymine to mispair with guanine and its mutagenic effect was originally ascribed to this property (1).Although other mechanisms are known to operate (Reviewed in (2), it has been conclusively demonstrated that the ^{Br}U.G mismatch is a major source of bromouracil induced transition mutations (2,3,4,5).

Recent experiments with synthetic templates and T4DNA polymerase (4) have shown that d^{Br}UMP is misinserted opposite G at a frequency of about 1% in competition with the correct substrate dCMP. Proofreading removes bromouracil with an efficiency of 82%, leading to an overall misincorporation rate of 0.16%. Under similar conditions, misinsertion of thymine is too small to be measured. Despite intense interest in the mutagenic

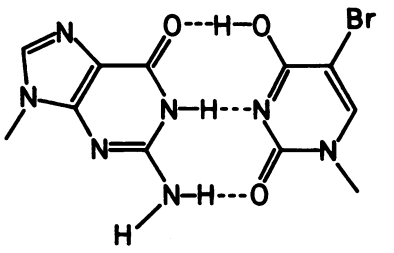
properties of 5-bromouracil, the exact nature of the ^{Br}U.G mismatch which gives rise to transition mutations has not yet been established. Three mechanisms have been put forward: tautomer shift, ionisation and base stacking.

The first model was originally suggested by Watson and Crick (6) who proposed that spontaneous mutations might be due to the occasional occurrence of a purine-pyrimidine pair with a base in one of its minor tautomeric forms. In this scheme, the rare enol form of thymine forms a three hydrogen bond base pair with guanine, but can not pair with adenine. Figure 1a illustrates the putative ^{Br}U(enol).G base pair. Spectroscopic investigations (7) indicate that some halogen substituents at the 5 position of uracil shift the keto- enol equilibrium towards the enol side, an observation which has been used to explain the mutagenic properties of 5-bromouracil(8). Because of the low abundance of the enol form the tautomer model is difficult to test.

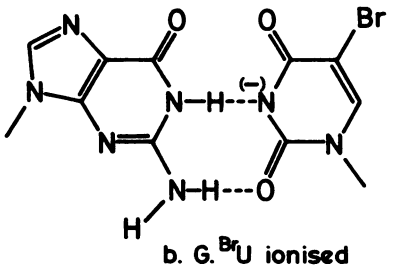
According to the ionisation model (Figure.1b) the electronegative bromine substituent induces ionisation at the N-3 position of uracil, thereby producing a configuration that is more likely to pair with guanine than adenine (9). Indeed the pKa value at the N-3 of the model compound 1-methyl 5-bromouracil is 7.8, so at physiological pH a significant proportion of ionised species is present. In contrast the corresponding pKa for 2'-deoxythymidine is 9.9.

In the base stacking model (10) the proposed ^{Br}U.G mismatch adopts the wobble configuration (Figure.1c). G.T wobble base pairs have recently been observed in a number of deoxyoligonucleotide crystal structures analysed in our laboratory(11,12). It has been suggested that ^{Br}U, when incorporated into DNA, forms stronger base stacking interactions than thymine and may thus give rise to more stable mismatches(10). The close contacts between halogen atoms and pyrimidine rings of neighbouring bases observed in the crystal structure of 5-bromouracil(10) give support to this model.

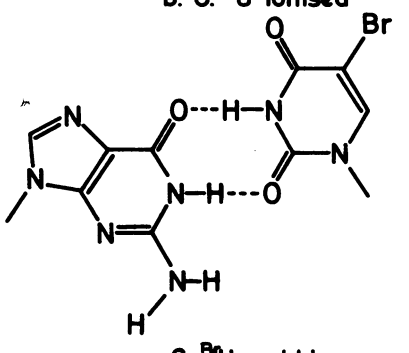
As part of a programme of investigating DNA fragments containing non complementary bases and other modifications, and in order to characterise the 5^{Br}U.G mismatch we synthesised the oligomer



a. G.^{Br}U enol



b. G.^{Br}U ionised



c. G.^{Br}U wobble

Figure 1 Schematic diagrams of proposed structures for the BrU.G base pair (a) Tautomer base pair (b) Ionised base pair (c) Wobble base pair.

d(^{Br}U-G-C-G-C-G), a self complementary hexamer which forms a duplex with ^{Br}U.G mismatches as terminal base pairs. The sequence was chosen because of its similarity to the Z-DNA hexamer d(C-G-C-G-C-G) studied by Rich and co workers at 0.9Å resolution (13) and which yielded by far the most detailed structure of any DNA fragment yet examined. It was necessary in the present case to aim for high resolution to detect a mixture of different base pair configurations at the mismatch site.

EXPERIMENTAL

Crystallisation and data collection

The hexanucleotide was synthesised by solid phase triester methods (14) and purified by ion exchange chromatography and hplc. It was crystallised above pH 8 so as to increase the proportion of the N-3 ionised form without significantly ionising N-2 of G (pK 9.8). A 14 µl droplet containing DNA (2 mM duplex), magnesium chloride (20 mM) and spermine tetrahydrochloride (2mM) all made up in a mixture of TRIS hydrochloride (50mM) and sodium acetate (25mM) buffered to pH 8.4 was placed in a crystallisation well and cooled from 18°C to 4°C. Vapour diffusion from an external well of 5% isopropanol in water produced crystals which attained a maximum size of 0.1 x 0.1 x 0.2 mm after two weeks. These crystals were too small to diffract to high resolution. Numerous attempts to increase their size were unsuccessful. Moreover, at higher pH, it was not even possible to grow crystals of this size.

The pH of a solution containing all the components of the crystallising mixture, except DNA, and with 5% isopropanol, was measured on an accurate pH meter and found to be 8.2.

Crystals of d(^{Br}U-G-C-G-C-G) were isomorphous with the parent Z DNA hexamer (13) and with d(T-G-C-G-C-G), a hexamer containing two G.T mismatches, which is currently being analysed at a resolution of 1.5Å (15). The space group is orthorhombic $P2_12_12_1$, a=17.94, b=30.85, c=49.95 A with two strands (or one duplex) in the asymmetric unit. Intensities were measured at 4°C on a Syntex $P2_1$ diffractometer. One crystal, housed in a sealed capillary tube, was used to collect three symmetry related data sets. These were corrected for the usual factors, including time dependent decay (max. 20%).After merging 1336 independent reflexions were obtained, to a resolution of 2.25Å. Of these, 1105 had $F>2\sigma(F)$.

Structure refinement

A model of d(^{Br}U-G-C-G-C-G) was constructed, using the coordinates of the parent compound d(C-G-C-G-C-G) but with ^{Br}U substituted for C. The model was first refined as a rigid body using the SHELX program (16) and then by the constrained-restrained program CORELS(17), where the sugars, the

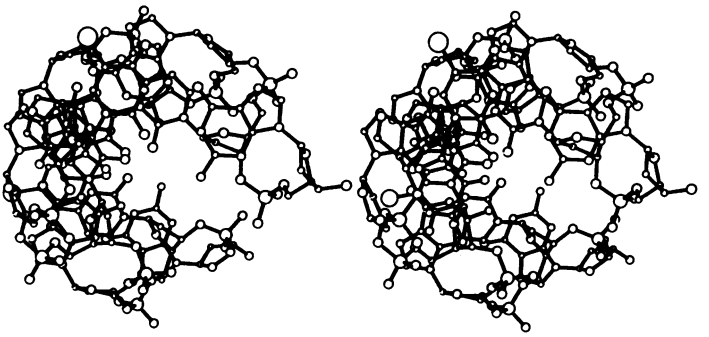


Figure 2 A stereoview of the double helix in d(^{Br}U-G-C-G-C-G), above the helix axis.

bases and the phosphates were treated as rigid groups. On convergence (R=0.27) an Fo-Fc map was calculated and solvent molecules were located. The refinement was continued with the Hendrickson-Konnert program(18) and the gradual inclusion of further solvent molecules, identified from Fo and Fo-Fc maps. Individual isotropic thermal parameters were included as variables. The refinement converged at R=0.156, with 64 solvents molecules (treated as water oxygens) located.

RESULTS AND DISCUSSION

The refined structure, illustrated in Figure 2 is in the left handed Z-DNA conformation, with dinucleotide repeat units. The average torsion angles are given in Table 1. They do not differ significanty from average values of Z-DNA hexamers, containing complementary base pairs only. The glycosidic torsion angle δ is anti for pyrimidines (-144^{o}) and syn for purines ($+71^{o}$). Similarly the torsional angle defining the sugar ring pucker χ alternates between 142^{o} for pyrimidines and 89^{o} for the purines. The average helical twist is 62.0^{o} and the rise 7.4Å between the repeat units. The backbone adopts the Z_{I} conformation except for the second GpC step of one strand, which is in the Z_{II} conformation. Such variations have been observed in other Z hexamer structures, and are dependent on the solvent environment around the phosphate groups (19). There are no clear differences in the local torsional angles around the normal Watson Crick and the mismatched base pairs. It appears that the

Table 1

Averaged Helical Parameters for Z-DNA Structures

	twist/dimer (°)	dimer/turn	rise/dimer (A)	base tilt (°)	Propellor twist (°)
CGCGCG	60	6	7.6	-6.2	4.6
TGCGCG	61.4	5.9	7.6	-8.2	5.4
^{Br}UGCGCG	62.0	5.8	7.4	-4.9	6.39

Average Torsional Angles for Z-DNA Structures

	α		β		γ		δ		ε		ς		χ	
CGCGCG	-137	48	-139	179	56	-170	138	100	-94	-104	80	-69	-150	72
TGCGCG	-149	71	-138	177	65	179	138	101	-92	-114	74	-56	-151	65
^{Br}UGCGCG	-165ǂ	105	-127	-175	57	155	142	89*	-83	-115ǂ	67	71ǂ	-144	71

Main chain torsional angles are defined by:

$$P\text{-}^{\alpha}O5'\text{-}^{\beta}C5'\text{-}^{\gamma}C4'\text{-}^{\delta}C2'\text{-}^{\varepsilon}O3'\text{-}^{\varsigma}P$$

The first figure in each column refers to a pyrimidine and the second to a purine pair.

* Not including the 3' terminal residues, G6 and G12
ǂ Not including the GpC step in the Z_{II} conformation

inclusion of two ^{Br}U.G wobble base pairs does not greatly perturb the global conformation of the double helix. Equally good fit of G.T wobble base pairs was observed in another Z hexamer (12,15) in two A-DNA octamers, (11,19) and in a B-DNA dodecamer (12). Full details of the individual torsional angles, base stacking and solvent structure for the hexamer will be presented elsewhere.

In the present communication we confine the discusssion to the nature of the mispairing at the mismatch sites. The ^{Br}U.G base pair, illustrated in Figure 3, is a typical wobble pair, similar in structure to the G.U pair observed in several tRNA crystal structures (20,21) and the G.T pair found in various oligonucleotides both in the solid state (9,10,) and in nmr solution studies (22).

Hydrogen bonds of length 2.7Å and 2.9Å (averaged over the two mismatched bases) connect N-1 of G with O-2 of ^{Br}U and the N-3 of ^{Br}U with O-6 of G respectively. There is a substantial shift of the bases from their position in a standard Watson-Crick pair. The displacement of the ^{Br}U into the major groove resembles that of thymine in the G.T mismatched stuctures referred to above.There is no indication of any disorder at the mismatch sites in the electron density maps calculated at the

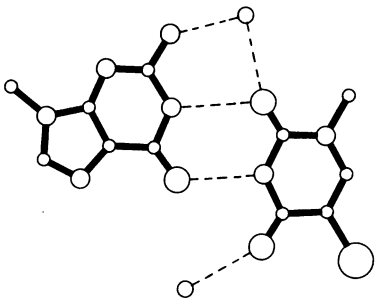


Figure 3 $^{Br}U_{Br}G$ wobble base pair observed in the crystal structure of d(^{Br}U-C-G-C-G-C). Note the two water molecules hydrogen bonded to the functional groups of the bases.

end of the refinement, and examined on a PS 300 graphics system. Thus the tautomer (Figure 1a) or the ionised (Figure 1b) pair could only be present in extremely small proportions.

The fact that the ionised base pair could not be detected in this specific structure, despite the high pH of the crystallisation mixture, suggests that the pKa of the N-3 atom of ^{Br}U is substantially increased when this base is incorporated into a DNA helix. The instability of the ionised pair may be due to the close proximity of two hydrogen acceptors,the O-4 of bromouracil and the O-6 of guanine,(approx 2.8Å) that results from such a rearrangement. Neverthless, we cannot entirely discount the possibility that either the tautomer or the ionised form of bromouracil is involved during DNA synthesis, since only 1% of the base need to be present in this form to account for its high frequency of misinsertion.

If the ^{Br}U.G wobble pair formation is a valid model for incorporation into DNA leading to transition mutations, what are the factors which characterise its mutagenic properties? Since the ^{Br}U.G and T.G wobble base pairs are essentially isosteric, increased incorporation of bromouracil is unlikely to result from structural differences between the two mismatches. Once incorporated,however, the ^{Br}U.G base pair may be more stable and escape detection more frequently. There are indications in support of this suggestion since DNA duplexes containing BrU.G base pairs have higher Tm values than the corresponding helices

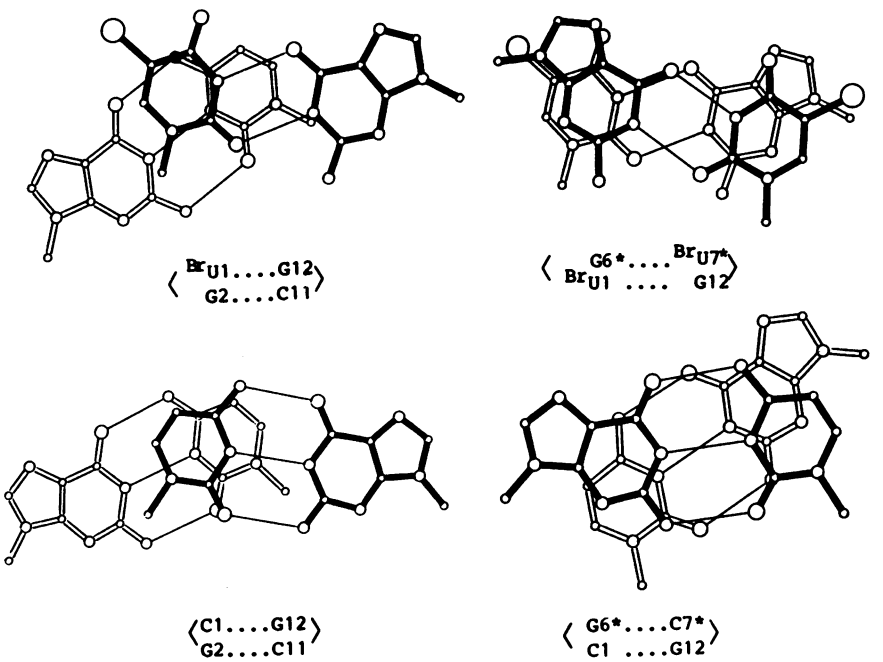


Figure 4 Comparison of base stacking in the hexamer d(^{Br}U-G-C-G-C-G) with the equivalent steps in d(C-G-C-G-C-G)(13). The upper diagram shows the intrastrand stacking of the ^{Br}U1.G12 base pair on the G2.C11 pair (left), in the 5'to3' direction and the interstrand stacking of the terminal G6.^{Br}U7 basepair of a symmetry related molecule (right). The lower two diagrams show the same intra and inter strand stacking at the corresponding steps in an idealised d(C-G-C-G-C-G) helix (Wang, A.H.-J. Protein Data Bank, Brookhaven). The view direction is perpendicular to the best plane betwen the two base pairs. Solid lines indicate the base pair nearest to the viewer.

with T.G base pairs (23). As can be seen there is indeed improvement of stacking, compared with the native hexamer (13). This could be a result of more favourable stacking interactions or more stable inter-base hydrogen bonds. Fig 4 illustrates the stacking of the ^{Br}U.G base pair on the adjacent G.C pair in the same helix in the 3' direction and the intermolecular stacking with the ^{Br}U.G base pair in the 5'direction. We are currently synthesising a number of deoxyoligonucleotides containing different base pair mismatches to test further this hypothesis, and are refining several related mismatched structures. These

will be compared in detail with the present structure in future publications.

ACKNOWLEDGEMENTS

We thank Steve Salisbury for preparing the monomer and for many fruitful discussions, Naveen Anand for helpful comments and the MRC Laboratory for Molecular Biology for access to the graphics facility. Financial support was provided by MRC under a Program Grant (to OK).

REFERENCES

1. Freese,E., (1950) J.Mol.Biol. 1,87-105.
2. Hopkins,R. and Goodman,M.F., (1980)Proc.Nat.Acad.Sci.USA,77 1801-1805.
3. Skopek,T.R. and Hutchinson,F., (1982) J.Molec.Biol.159, 19-33.
4. Lasken,R.S. and Goodman,M.F., (1984) J.Biol.Chem. 259, 11491-11495.
5. Lasken,R.S. and Goodman,M.F., (1985) Proc.Nat.Acad.Sci.82, 1301-1305.
6. Watson,J.D. and Crick,F.H.C., (1953) Nature 171, 737-738.
7. Katritzky,A.R. and Waring,A.J. (1962) J.Chem.Soc.1540-1544.
8. Topal,M.D. and Fresco,J.R., (1976) Nature 263,285-289.
9. Lawley, P.D. and Brookes, P. (1962) J.Molec.Biol.4, 216-219
10. Sternglantz,H. and Bugg,C.E., (1975) Biochim. Biophys.Acta 378, 1-11
11. Brown. T., Kennard, O.,Kneale.G.., and Rabinovich, D., (1985) Nature 315 604-606.
12. Kennard,O., (1985) J.Biomol.Str.Dyns.3, 205-226.
13. Wang,A.H.J., Quigley,G.J., Kolpack,F.J., Crawford,L., van Boom,J.H., van der Marel,G., Rich,A., (1979) Nature (London) 282 680-686.
14. Gait,M.J., Matthes,H.W.D., Singh,M., Sproat,B.S and Titmus,R.C. (1982) Nucleic Acid Res. 10 6243-6248.
15. Rabinovich,D., Brown,T., Kneale,G.G., Kennard,O., to be published.
16. Sheldrick,G.M., (1976) SHELX 76 System of Computing Programs, University of Cambridge, England.
17. Sussman,J.L., Holbrook,S.R., Church,G.M. and Kim,S.-H.,(1977) Acta Cryst. 33, 800-804.
18. Hendrickson,W.A. and Konnert,J.H., (1981) in Biomolecular Structure, Conformation, Function and Evolution, Vol.1,(ed. Srinivasan,R.) 43-57, Pergamon, Oxford.
19. Kneale,G.G., Brown,T., Kennard,O. and Rabinovich,D., (1985) J.Molec.Biol. (in press)
20. Jack,A., Ladner,J.E. and Klug,A., (1976) J.Molec.Biol. 108,619-649.
21. Rich,A., (1977) Acc.Chem.Res. 10,388-402.
22. Patel,D.J., Kozlowski,S.A., Ikuta,S. and Itakura,K., (1984) Fedn.Proc. 43,2663-2670.
23. Anand,N., (unpublished data)