Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Exploration of the influence of 5-iodo-2′-deoxyuridine incorporation on the structure of d[CACG(IDU)G]

Geertrui Schuerman,[a,b] Kristof Van Hecke[a] and Luc Van Meervelt[a]*

[a]Biomoleculaire Architectuur, Departement Chemie, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001 Leuven (Heverlee), Belgium, and [b]Vlaams Interuniversitair Instituut voor Biotechnologie (VIB), Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel, Belgium

Correspondence e-mail: luc.vanmeervelt@chem.kuleuven.ac.be

Received 12 March 2003
Accepted 4 June 2003

PDB Reference:
d[CACG(IDU)G], 1omk, r1omksf.

The first antiviral nucleoside 5-iodo-2′-deoxyuridine (IDU) against herpes simplex virus type 1 and type 2 is a thymidine analogue, *i.e.* the C5 methyl group is replaced by an I atom. The structure of the self-complementary hexamer d[CACG(IDU)G] was determined by single-crystal X-ray diffraction techniques. The orthorhombic crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 18.16, b = 30.03, c = 41.99 Å. Refinement in the resolution range 20–1.3 Å converged with a final $R1$ = 0.167, including 43 water molecules and two cobalt hexammine complexes. The incorporation of a large I atom has only minor consequences for the overall structure as is noticed in the IDU·A base pairs, which are of the common Watson–Crick type. To contribute to the still puzzling mechanism of this historically important agent, details of base stacking, helical parameters, hydration *etc*. have been studied. A general scheme of cobalt hexammine-binding modes in Z-DNA is provided, revealing similar binding modes for the reported structure.

1. Introduction

The discovery of the first selective antiviral nucleoside 5-iodo-2′-deoxyuridine (IDU) (Prusoff, 1959; Kaufmann, 1962), specifically against herpes simplex virus type 1 (HSV-1), has led to a broad spectrum of other 5-substituted 2′-deoxyuridine compounds and antiviral drugs such as acyclovir and ganciclovir (De Clercq *et al.*, 1979). Today, IDU is still used in the treatment of herpetic keratitis, mucocutaneous infections and for tumour radiosensitization in patients with high-grade gliomas (Kinsella *et al.*, 2000).

The target enzyme in the activation of the modified nucleoside is the virus-encoded thymidine kinase (TK), which phosphorylates thymidine to its corresponding 5′-monophosphate. The latter is then phosphorylated by cellular kinases to its corresponding thymidine-5′-triphosphate, which is the precursor of the thymidine incorporation in DNA. As viral TK has a much broader substrate specificity than the corresponding enzyme in humans (Spadari *et al.*, 1992), phosphorylation occurs preferentially by the viral TK and the modified nucleoside is activated only in the infected cells. However, the antiviral activity can only be explained at the level of some other crucial viral enzymes such as DNA polymerase, reverse transcriptase *etc*.

Considering the historical importance of IDU and the still unravelled mechanism of interaction, a contribution is made by the structure determination of the sequence d[CACG(IDU)G] by single-crystal X-ray diffraction techniques. The reported 1.3 Å structure has the left-handed Z-conformation and contains two ordered cobalt hexammine complexes. Comparing several cobalt hexammine-containing structures, some general patterns can be detected in the binding of those complexes to different residues and different duplexes.

2. Materials and methods

2.1. Crystallization and data collection

Synthesis and purification of the hexanucleotide d[CACG(IDU)G] were performed at the Rega-Institute of the K.U. Leuven. Crystals were grown in hanging drops by vapour diffusion from a 4 µl drop consisting of 2 µl crystallization solution and 2 µl of a 7.66 mg ml^{-1} solution of d[CACG(IDU)G] in 28.7 m*M* potassium cacodylate buffer pH 6.9 equilibrated against 500 µl 24.5%(*v*/*v*) MPD (4-methyl-2,4-pentanediol) reservoir solution. The crystallization solution contained 40 m*M* potassium cacodylate buffer pH 5.5, 6 m*M* cobalt hexammine, 80 m*M* KCl, 20 m*M* $MgCl_2$, 10%(*v*/*v*) MPD. The oligonucleotide crystallized in space group $P2_12_12_1$, with unit-cell parameters a = 18.155, b = 30.034, c = 41.988 Å. The asymmetric unit contains one double-stranded DNA helix.

One single-crystal (0.25 × 0.25 × 0.15 mm) was used to collect a 98.2% complete data set at EMBL beamline X11 of the DESY in Hamburg. The data were collected at 120 K and a wavelength of 1.000 Å using an 18 cm

Table 1
Data-collection statistics for the structure d[CACG(IDU)G].

Values in parentheses are for the outermost resolution shell.

Resolution range (Å)	20.0–1.30 (1.35–1.30)
Measured reflections	137352
Unique reflections	5988
Data-to-parameter ratio	5.0
Completeness (%)	98.2 (98.8)
R_{sym}	6.2 (16.5)
Multiplicity	22.9
Mean $I/\sigma(I)$	6.3
Reflections with $I > 3\sigma(I)$ (%)	78.1 (45.9)

MAR Research image plate, yielding 5988 reflections in the resolution range 20–1.3 Å [of which 84.4% have $I > 2\sigma(I)$], with R_{sym} = 6.2%. Data were processed with *DENZO* (Otwinowski & Minor, 1997) including overloads, and reduced using *SCALEPACK* (Otwinowski & Minor, 1997). Data-collection statistics are given in Table 1.

2.2. Structure solution and refinement

A molecular-replacement solution was found using *AMoRe* (Navaza, 1994) and the structure d[CGCGC(Br)G] (NDB code ZDFB51) as a starting model. The structure was refined using *SHELXL*97 (Sheldrick & Schneider, 1997). The structure contains two cobalt hexammine molecules. After the addition of 43 water molecules, both I atoms were refined anisotropically, converging to a final ωR_2 value of 42.6% and an R value of 16.7% for all the data.

The nucleotides of strand 1 are labelled C1–G6 in the 5′ to 3′ direction and C7–G12 on strand 2.

3. Results and discussion

3.1. Overall helix structure

The self-complementary hexamer d[CACG(IDU)G] adopts the double helical Z-DNA conformation. Both A·IDU pairs are of the Watson–Crick type, indicating only minor consequences for the overall structure. When fitting the reported structure to the most similar structure $d[C(NH_2)ACGTG]_2$ (NDB code ZDFB11) from Coll *et al.* (1986), an r.m.s. deviation value of 0.78 Å is found, confirming the findings above.

The helical parameters as calculated by the program 3*DNA* (Lu *et al.*, 2000) are typical for Z-DNA, with an average helical twist of −57.4° for each dinucleotide repeat.

Backbone torsion angles are illustrated in Table 2, most of the torsion angles fall well within the ranges typical of Z-DNA.

Table 2
Backbone torsion angles, glycosidic torsion angles and pseudorotation angles P as calculated by the program 3*DNA* (Lu *et al.*, 2000) for the structure d[CACG(IDU)G)].

Backbone torsion angles are defined as: O3′(i − 1)–P–α–O5′–β–C5′–γ–C4′–δ–C3′–ε–O3′–ξ–P(i + 1)–O5′(i + 1). Glycosidic torsion angles χ for pyrimidines (Y) O4′–C1′–N1–C2 and for purines (R) O4′–C1′–N9–C4.

	α (°)	β (°)	γ (°)	δ (°)	ε (°)	ξ (°)	χ (°)	P (°)
C1			50.8	134.3	−97.3	80.7	−150.3	151.6
A2	56.1	−166.0	−176.8	94.8	−178.0	55.1	67.5	39.6
C3	170.1	152.7	53.9	143.9	−99.4	79.8	−154.9	152.4
G4	67.0	−165.9	179.1	96.2	−127.8	−38.6	58.7	35.0
IDU5	−159.0	−139.3	49.3	138.1	−98.2	66.5	−152.9	138.5
G6	82.4	−174.6	−173.9	144.6			75.8	166.9
C7			49.3	138.5	−88.1	86.3	−153.3	153.8
A8	−156.7	174.6	47.8	77.6	−156.1	−57.3	80.4	52.0
C9	−152.3	−117.1	41.9	137.3	−93.8	63.0	−158.8	140.6
G10	73.5	−175.7	−176.0	99.1	−138.7	−9.1	60.6	54.6
IDU11	−170.0	−171.2	55.1	139.1	−101.9	57.6	−158.9	152.8
G12	84.4	175.3	−164.4	163.1			74.0	172.8

The pyrimidine (including IDU) nucleosides have *anti* glycosidic torsion angles (in the range −158.9 to −150.3°), with C2′-*endo* sugar puckering tending to C1′-*exo* (for IDU5 and C9). All purine nucleosides have *syn* glycosidic torsion angles (in the range 58.7–80.4°) that are a little larger than those typical for Z-DNA. They have C4′-*exo* sugar puckers, except for G4 which has a C3′-*endo* sugar pucker. G6 and G12 have C2′-*endo* puckers which are not unusual for end-standing guanosines.

The unusual value for torsion angle α of adenine A8 (−156.7°) results in the protruding of C5′ (A8) out of the duplex (Fig. 1).

The less common Z_{II} conformation has previously been found mostly at P5 in contrast with the reported structure where P3 is in the Z_{II} conformation. The Z_{II} conformation has also been observed at P3 in two other Z-DNA hexamers (Schroth *et al.*, 1993; Harper *et al.*, 1998).

In general, the Z_{II} conformation is stabilized by nearby polyvalent cations. However, also in the absence of cations the Z_{II} conformation has been observed in the so-called 'pure spermine' form (Egli *et al.*, 1991).

In the reported structure, the Z_{II} conformation is slightly stabilized by the nearest cobalt hexammine ion (Co30 of a neighbouring duplex) with O1P(C3)···N distance of 3.46 Å.

3.2. The A·IDU Watson–Crick base pairs

Both A·IDU base pairs are of the Watson–Crick type with hydrogen-bond distances between 2.68 and 2.93 Å. The A·IDU base pairs are virtual identical, with an r.m.s. deviation of only 0.25 and 0.45 Å

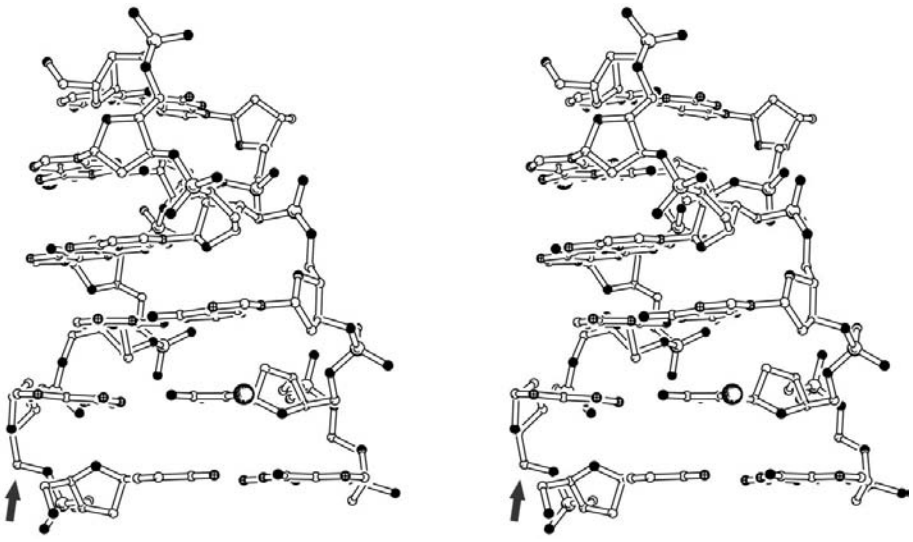

Figure 1
Stereoview in stick representation of the hexamer d[CACG(IDU)G]. Base pair C1·G12 is on top. The position of C5′ (A8) is indicated. Picture generated using *PLATON* (Spek, 2000)

for the fitted bases and bases with sugars, respectively.

The thermal vibration of base pair A2·IDU11 is higher compared with that of IDU5·A8 (data not shown). This difference can be explained by a short O2P(G4)···O2P(G12) distance (5.03 Å) caused by cobalt hexammine Co31 bridging phosphates 4 and 12. The shortening by Co31 results in an impaired stacking of O4′(IDU11) with G12 and of the I atom of IDU11 with G10. Finally, the A2·IDU11 pair is also less efficiently hydrated compared with IDU5·A8.

3.3. Interactions with cobalt hexammine

Searching the NDB for Z-DNA sequences containing a cobalt hexammine complex in the asymmetric unit resulted in seven structures. In all these structures cobalt hexammine exhibits a similar intramolecular binding mode, *i.e.* it binds to nitrogen N7 and oxygen O6 of a guanine residue, and to a phosphate oxygen of the subsequent phosphate of the same duplex (shown in red in Fig. 2*a*). Beside this intramolecular binding mode, the cobalt hexammine is in contact with phosphate O atoms of one or two phosphates of a symmetry-equivalent duplex (Fig. 2*a* in red and yellow). Two of these seven structures contain a second cobalt hexammine complex: ZDF060 and ZDG057. In ZDF060 the second cobalt hexammine does not make contact with a base moiety but only contacts two phosphates of the same duplex (Fig. 2*b*). In ZDG057, the second cobalt hexammine undergoes intramolecular binding but without making contact to the guanine O6, and makes contacts to phosphates belonging to different symmetry-equivalent duplexes (shown in red, yellow and green in Fig. 2*a*).

In the present structure these general schemes vary slightly. The cobalt 31 complex shows the intramolecular binding mode, as contacts with O6(G6) and N7(G6) are made, but without contact with any phosphate atom of the same duplex. However, the complex bridges phosphates O2P(G12) and O2P(G4) of a neighbouring symmetry-related duplex as in most of the cases. Thus, in general the scheme in Fig. 2(*a*) is followed. The cobalt 30 complex interacts with O2P(G10), O1P(G6) and O2P(G6) of the same duplex. Contacts with O2P(IDU5) and O3′(G4) of a symmetry-related duplex as well as with N7(G4) of another symmetry-related neighbouring duplex are made (Fig. 2*c*). These contacts are quite similar to those seen for the second cobalt hexammine in ZDF060 (Harper *et al.*, 1998). In our case the extra contacts that appear can be explained by the tighter packing.

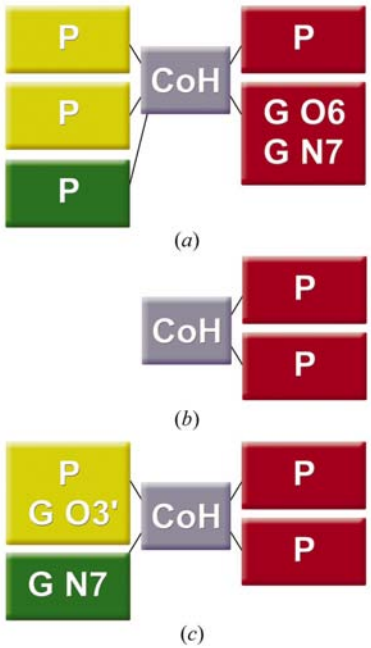

Figure 2
Schematic representation of the intramolecular and intermolecular binding modes of cobalt hexammine complexes to Z-DNA. Red coloured blocks represent atoms of the same asymmetric unit; yellow and green blocks represent atoms from different symmetry-related duplexes. Cobalt hexammine complexes are shown as grey blocks. (*a*) Intramolecular binding mode in red, additional phosphate contacts in yellow and green. (*b*) Binding mode of second cobalt hexammine complex in ZDF060. (*c*) Binding mode of complex 30 in the reported structure.

3.4. Packing

The crystal packing belongs to the so-called mode 2 or pure-spermine form, first described by Egli *et al.* (1991). The only other Z-DNA structure (containing no overhanging bases) crystallized with cobalt hexammine, d(TGCGCA)$_2$, also adopts this mode (Harper *et al.*, 1998). The reduction of the unit-cell volume when going from mode 1 (25036 Å^3) to mode 2 (24436 Å^3), and caused by a compression along the helical axis (Egli *et al.*, 1991), is even more pronounced in the reported structure (22 895 Å^3). This compact DNA conformation is also apparent from the P···P distances. The intrastrand P···P distances are on average slightly smaller than those in the pure-spermine form, but the interstrand P···P distances (8.00 Å averaged over the three distances) are much shorter than those in the pure-spermine form (9.12 Å averaged over the three distances), making the minor groove even more narrow compared with the pure-spermine form. It seems that in the structure under investigation the DNA is maximally compressed in the direction perpendicular to the helix axis. The reduction in unit-cell volume is caused by a compression in all three directions, which is a combined effect of crystal packing and data collection at 120 K.

4. Conclusions

In Z-DNA the relative orientation and position of the base pairs with respect to the helix axis are determined by the rigid structure of the backbone and thus are only slightly affected by the base sequence or iodination of one of the bases as in our case. This is illustrated in the structure of d[CACG(IDU)G] containing two A·IDU Watson–Crick base pairs that shows no unusual conformational features. Calculated parameters are in the range typical for Z-DNA. The structure contains the Z_I/Z_{II} conformational mix, with P3 adopting the Z_{II} conformation, and a cobalt hexammine ion in its neighbourhood (Co30 of a neighbouring duplex).

Although it contains two cobalt hexammine complexes in its asymmetric unit, the structure crystallizes in the pure spermine form or mode 2.

The intramolecular binding mode of cobalt hexammine complexes generally found is mainly preserved for one cobalt hexammine complex in the present structure. The second cobalt hexammine complex shows a similar binding mode as in ZDF060. The two cobalt hexammine complexes stabilize the packing by forming bridges connecting symmetry-related DNA helices rather than by stabilizing Z_{II} conformations.

The two A·IDU base pairs differ in hydration and stacking interactions, which is further reflected in a difference in thermal vibration. The differences in stacking interactions are caused by a cobalt hexammine complex, showing that such molecules, together with the iodination at the C5 position of a thymidine, play an important role in the stacking of residues. However, despite these differences, the geometry of both A·IDU base pairs is similar.

The hydrophobic character of the I atom makes IDU the ideal ligand of HSV-1 TK in the active site, which mostly consists of non-polar side chains (Champness *et al.*, 1998). In agreement, we observe only one water molecule within a radius of 3.5 Å (sum of the O and I van der Waals radii) around the two I atoms.

This work was supported by the Fund for Scientific Research (Flanders), K.U. Leuven and the European Community Access to Research Infrastructure Action of the Improving Human Potential Program to the EMBL Hamburg Outstation. We thank the staff of the EMBL Hamburg Outstation for their help with the synchrotron experiments, and P. Herdewijn, A. Van Aerschot and G. Schepers for synthesis of the oligonucleotide.

References

Champness, J. N., Bennett, M. S., Wien, F., Visse, R., Summers, W. C., Herdewijn, P., de Clercq, E., Ostrowski, T., Jarvest, R. L. & Sanderson, M. R. (1998). *Proteins*, **32**, 350–361.

Coll, M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H. & Rich, A. (1986). *J. Biomol. Struct. Dyn.* **4**, 157–172.

De Clercq, E., Descamps, J., De Somer, P., Barr, P. J., Jones, A. S. & Walker, R. T. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 2947–2951.

Egli, M., Williams, L. D., Gao, Q. & Rich, A. (1991). *Biochemistry*, **30**, 11388–11402.

Harper, A., Brannigan, J. A., Buck, M., Hewitt, L., Lewis, R. J., Moore, M. H. & Schneider, B. (1998). *Acta Cryst.* D**54**, 1273–1284.

Kaufmann, H. E. (1962). *Proc. Soc. Exp. Biol. Med.* **109**, 251–252.

Kinsella, T. J., Vielhuber, K. A., Kunugi, K. A., Schupp, J., Davis, T. W. & Sands, H. (2000). *Clin. Cancer Res.* **6**, 1468–1475.

Lu, X.-J., Shakked, Z. & Olson, W. K. (2000). *J. Mol. Biol.* **300**, 819–840.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Prusoff, W. H. (1959). *Biochim. Biophys. Acta*, **32**, 295–296.

Schroth, G. P., Kagawa, T. F. & Ho, P. S. (1993). *Biochemistry*, **32**, 13381–13392.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Spadari, S., Maga, G., Focher, F., Ciarrocchi, G., Manservigi, R., Arcamone, F., Capobianco, M., Carcuro, A., Colonna, F. & Iotti, S. (1992). *J. Med. Chem.* **35**, 4214–4220.

Spek, A. L. (2000). *PLATON*. Molecular Geometry Program. Utrecht University, Utrecht, The Netherlands.