Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Destabilizing effect of a fluorouracil extra base in a hybrid RNA duplex compared with bromo and chloro analogues

William Cruse,[a] Pedro Saludjian,[a] Alain Neuman[a] and Thierry Prangé[a,b]*

[a]CSSB – Chimie Structurale Biomoléculaire (UPRES A7031 CNRS), UFR Biomédicale 74, Rue Marcel Cachin, 93017 Bobigny CEDEX, France, and [b]LCRB – Laboratoire de Cristallographie et RMN Biologiques (UMR 8015, CNRS), Faculté de Pharmacie, 4 Avenue de l'Observatoire, 75270 Paris CEDEX 06, France

Correspondence e-mail: prange@lure.u-psud.fr

Received 2 April 2001
Accepted 19 July 2001

PDB References: r-GCUUCGGC-d^{Cl}U, Rh(NH_3), 1idw; r-GCUUCGGC-d^{Br}U, Rh(NH_3), 1iha; r-GCUUCGGC-d^{F}U, Rh(NH_3), 1id9; r-GCUUCGGC-d^{F}U, Ir(NH_3), 1icg.

In the presence of cobalt, rhodium or iridium hexammine salts, the RNA/DNA hybrid r-GCUUCGGC-d^{X}U (with X = F, Cl or Br) crystallizes as a double-stranded helix with four consecutive G–U and C–U mismatches. The deoxy chloro- and bromouracil derivatives are isomorphous, space group $C2$, unit-cell parameters $a = 53.80$, $b = 19.40$, $c = 50.31$ Å, $\beta = 109.9°$, with the same infinite helix arrangement in the packing along the c axis with one extra DNA halogenouracil base included in the stacking. However, the fluorouracil derivative, with unit-cell parameters $a = 53.75$, $b = 19.40$, $c = 45.84$ Å, $\beta = 105.7°$, is not isomorphous. The corresponding extra DNA base d^{F}U of the second strand is ejected out of the helical stack, leading to a shortening of the c axis. The specific destabilization of the fluorouracil for the duplex building is analyzed in terms of the polarization effect of the halogen atom attached to the 3′-terminal base that modulates its interactions.

1. Introduction

Although tetraloops or hairpin motifs in oligomers have been widely observed in solution by NMR (Davis *et al.*, 1993; Mirmira & Tinoco, 1996), the crystalline state usually favours the formation of duplexes (Holbrook *et al.*, 1991; Carter *et al.*, 1997). Accordingly, the X-ray structure of the non-complementary tetraloop motif r-UUCG embedded in a short complementary sequence revealed, instead of the expected hairpin loop, a duplex structure (Cruse *et al.*, 1994). In the present study, the RNA octamer r(GCUUGGC) was linked to different extra-DNA bases at the 3′ end. This halogenouracil was introduced in the hope of solving the structure by MAD (X = Br), to aid crystallization and/or to induce resistance against RNAses. Further, the hybrid nonamers crystallize only in the presence of cobalt, rhodium or iridium hexammine trichloride with an increasing resolution from 2.2 Å (cobalt hexammine) to ∼1.5 Å (iridium hexammine). Although these complexes are isomorphous whatever the hexammine cation in use, the fluorouracil hybrids are not isomorphous to the corresponding chloro- or bromouracil complexes.

A systematic investigation of the crystal structures of these complexes by varying either the hexammine cation or the extra base has been undertaken, with the aim of understanding the role of the extra fluorouracil compared with the chloro or bromo homologues.

2. Experimental data

The synthesis and purification of r(GCUCGGC)d^{Br}U has been described previously (Cruse *et al.*, 1994); the other nonamers r(GCUCGGC)-d^{X}U were also obtained from the

Table 1
Data reductions and refinement statistics.

All space groups are $C2$, $Z = 8$.

r-GCUUCGGC-d^{X}U, $M(NH_3)$	X = Cl/ M = Rh	X = Br/ M = Rh	X = F/ M = Rh	X = F/ M = Ir
PDB code	1idw	1iha	1id9	1icg
Unit-cell parameters				
a (Å)	53.82	53.80	53.73	53.75
b (Å)	19.38	19.40	19.39	19.40
c (Å)	50.30	50.31	45.83	45.849
β (°)	109.9	109.9	105.7	105.7
Volume per residue (Å^3)	686	685	637	639
No. of measured reflections	13027	40122	38535	42723
No. of unique reflections	3738	6400	4425	6772
Redundancy	3.5	6.2	9.0	6.5
Data completeness (%)	85	96	92	95
Overall $I/\sigma(I)$	18.8	21.0	21.0	15.1
Overall R_{merge} (%)	3.9	4.5	5.7	4.4
Resolution limits (Å)	6.0–1.8	9.5–1.57	11.0–1.6	10.0–1.53
No. of atoms in the model (including waters and ions)	440	452	436	440
No. of F_{obs} used in refinements [all/$F > 2\sigma(F)$]	3738/3641	6400/6312	4425/4339	6772/6518
R_{free} (%) (8% of data)	20.2	19.6	19.5	20.5
wR^2 ([$F_{obs} > 4\sigma(F^2)$])	0.443	0.391	0.476	0.414
R factors (%); $F_{obs} > 2\sigma(F)$/all	17.2/17.6	16.7/16.1	15.3/15.5	17.7/18.0

OSWEL DNA-synthesis service as the triethylamine salt and converted to ammonium salts using a strong cation-exchange resin. Rhodium (III) hexammine trichloride was prepared by heating 2 g $[Rh(NH_3)_5Cl]Cl_3$ in 200 ml concentrated ammonia for 36 h at 423 K in a sealed tube and subsequently purified by fractional crystallization from hot water. $[Ir(NH_3)_6]Cl_3$ was prepared the same way by heating 2 g K_3IrCl_6 with concentrated ammonia in a sealed tube for 48 h at 423 K and was purified by fractional crystallization until a colourless product was obtained.

The crystals were prepared by reverse vapour diffusion from sitting drops in well plates at 277 K. The initial crystallization liquor contained 50 µl of 0.05 *M* lithium cacodylate buffer pH 6.5, 0.1 mg RNA/DNA hybrid, 0.1 mg rhodium/iridium hexammine and 5% dimethyl pentanediol (MPD). The reservoir was 35–40% MPD solution. Monoclinic plates appeared after 24–48 h, growing to optimal dimensions of 0.2 × 0.6 × 0.05 mm over a period of three weeks. The crystallization process was reversible. Unsatisfactory preparations were redissolved by adding 50–100 µl of water and heating the reservoir at 313 K until solution was complete. Recooling and slow evaporation of the redissolved material produced crystals of diffraction quality. This process enabled complete structure determination with less than 4 mg of purified nucleotide in each case.

2.1. Data collection and processing

Data from the crystals of the bromo and fluoro derivatives were recorded at 278 K in wet capillaries at the wiggler line W32, LURE facility, Orsay (Fourme *et al.*, 1992) with a MAR 345 image-plate system at a wavelength λ = 0.967 Å. The chloro derivative was recorded on a FAST Nonius apparatus (λ = 1.5418 Å) at the limited resolution of 1.9 Å. The data sets are summarized in Table 1. In all cases, the temperature was 277 K.

Processing of the data was conducted with the *MOSFLM* program (Leslie, 1994) interfaced with the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994).

2.2. Structure determination

A model of the central region $(UUCG)_2$ was first built from the mismatch duplex described by Holbrook *et al.* (1991). The structure of the rhodium (III) hexammine–bromouracil complex was first solved with the many-body version of *AMoRe* (Navaza *et al.*, 1993, 1998) by dividing the whole structure into four pieces: two Watson–Crick duplexes $(G–C)_2$, the central $(UUCG)_2$ and the rhodium (III) hexammine cation. The oriented/translated fragments were reconstructed and assembled altogether in the same asymmetric unit on a graphic display. The construction was completed by a rigid-body refinement. At this stage (R = 39%), it was possible to build one of the extra DNA bases in the residual density at one end of the first strand but not at the other end, where the electron density of $F_o - F_c$ maps clearly indicated disorder. This last base was not included in the first rounds of refinement, which ended at R = 27%.

The non-isomorphous fluorouracil hybrid structures with M = Rh and Ir were obtained by MR using the central double helix of the bromo analogue without the extra DNA bases.

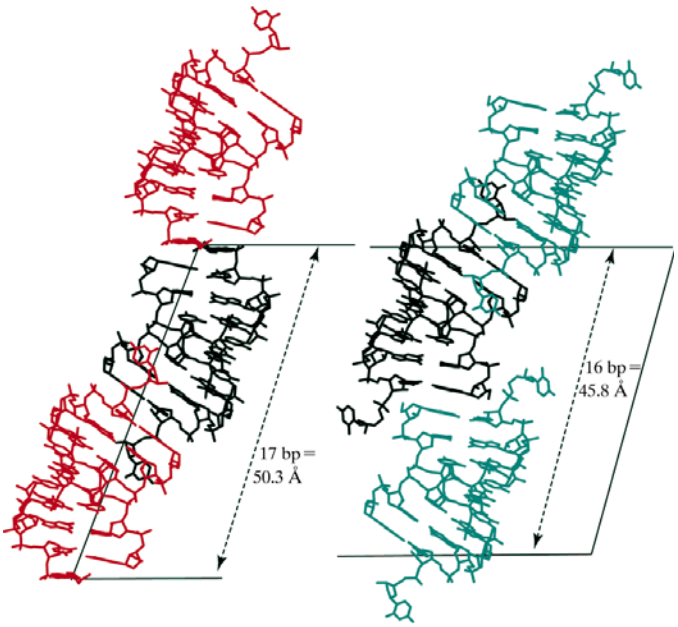

Figure 1
Infinite helices for (*a*) the bromouracil duplex (black/red) and (*b*) the corresponding fluorouracil-containing structure (black/blue). In the Br-U (and Cl-U) structures, the extra base 18 is part of the infinite helix built along the *c* axis, while in the case of the fluorouracil the corresponding base is rejected out of the stack. The *c* axis accounts for 16 base pairs in the fluoro (right) and for 17 in the bromo and chloro derivatives (left), including the intermolecular stacking; the average distances for the base pairs are 2.86 and 2.95 Å, respectively.

The two extra d-^{F}U bases were subsequently reconstructed in difference Fourier maps.

2.3. Refinements

The four rhodium or iridium hexammine complexes (Table 1) were refined using the *SHELXL* program (Sheldrick, 1997) at their nominal resolutions. The refinements were performed with standard restraints on bond (1–2) and angle (1–3) distances, planes and chiral (zero and non-zero) volumes. In the final stages, the base ^{X}U18 (X = Cl and Br) close to the twofold axis was modelled in two positions (1:1 ratio). This disordered base ^{X}U18 forms a U–U pair through the twofold axis of the space group with its symmetry-related mate, the first position (a) being connected to the symmetry-related second position (b*) and *vice versa* (Fig. 2*c*).

In the final round of refinement the Rh and/or Ir atoms were refined anisotropically. The r.m.s. deviations from ideality and statistics are reported in Table 1.

3. Results

Although cobalt hexammine cations have been cited as being useful for the crystallization of synthetic DNA fragments (Nunn, 1996) or larger t-RNA molecules (Hingerty *et al.*, 1982), there is a general assumption that they favour structures belonging to the Z-DNA family of helices (Geissner *et al.*, 1985; Ho *et al.*, 1987). In this paper, we report the structures of four synthetic RNA oligonucleotide nonamers which crystallize as duplexes in the A-form with cobalt, rhodium or iridium (III) hexammine cations. Cobalt (III) hexammine derivatives were obtained as yellowish crystals, but with a limited diffraction of 2.2 Å. Their structures were solved and found to be completely isomorphous to the rhodium (III) or iridium (III) hexammine complexes. Cobalt (III) complexes were not fully refined and are not reported here.

Each structure crystallizes as an antiparallel duplex (form *A*). As the double helix runs parallel to the *z* axis of the crystal lattice, the packing of symmetry-related duplexes on top of each other leads to the formation of continuous double-helical stacks in the crystal (Fig. 1). This style of packing is well documented in A-DNA and B-DNA structures in both the

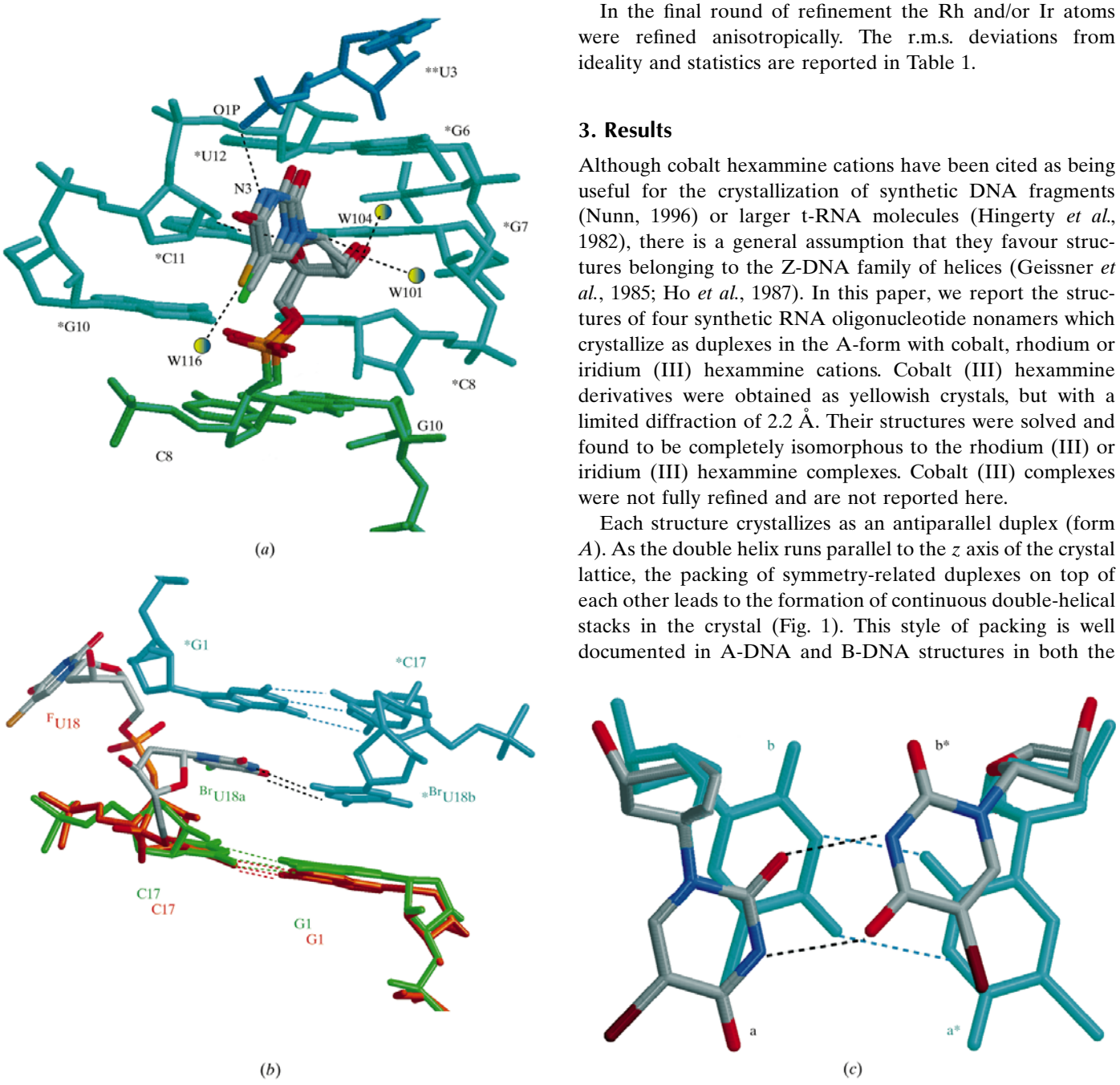

Figure 2
(*a*) The DNA halogenouracil 9 (end of the first strand) is ejected into a small cavity built by the packing. This residue is tightly bound by numerous hydrogen bonds: N3 to O1P of a symmetry-related U3 (d = 2.70 Å), O4′ to O2′ of **C11 (d = 2.68 Å), O3′ to N2 of **G7 (d = 2.98 Å) plus three additional water molecules (W101, W104 and W116). The three ^{F}U,^{Cl}U and ^{Br}U superimposed by least-squares calculations show the same orientation. (*b*) Least-squares fitting of the fluoro derivative (in orange) over the bromouracil chain (green, blue for the symmetry-related) at the 3′ end of the second strand. The halogenouracil 18 is either rejected out of the stack (X = F) or included in the stack (X = Cl, Br). For clarity, only one disordered position of the stacked bromouracil is shown. (*c*) View of the 1:1 disordered ^{Br}U–^{Br}U (or (^{Cl}U–^{Cl}U) pair at the *C*2 axis position. The first disordered base (a) is connected to the second of the symmetry-related duplex (b*) and *vice versa* (b connected to a*).

free and intercalated states (Gao *et al.*, 1990; Timsit & Moras, 1994; Cruse *et al.*, 1994, 1996; Carter *et al.*, 1997) and is considered to be an important stabilizing effect in the building and stability of crystals.

In each structure there are eight base pairs and one unpaired 5-halogeno deoxyuridine at each end. The full *c* axis of the cell accounts for 16 base pairs in a helical stack for the fluoro analogue and 17 base pairs for the bromo and chloro analogues, which are non-isomorphous to the fluoro analogue (two duplexes on top of each other in each case).

The stacking at one end of the helix is strengthened by the position of the sugar of the extra DNA base in the minor groove of the twofold symmetry-related helix. The dangling uracil base is hydrogen bonded to the phosphate backbone of a translation-related duplex and also makes a number of strong hydrogen bonds with symmetry-related strands and water molecules (Fig. 2*a*). The same situation holds for the chloro and the bromo analogues, all of which show clear electron densities in this region and equivalent orientations.

At the other end of the duplex, interactions maintaining the extra halogenouracil 18 are considerably weaker. Though also ejected from the helical stack, the ^{F}U18 resides in a large pocket delimited by the crystal packing with either disordered orientations or a high thermal motion. In contrast, the bromo and chloro analogues are not ejected but are an integral part of the helical stack at the position of the twofold axis, showing a disordered ^{Br}U–^{Br}U* (or ^{Cl}U–^{Cl}U*) intermolecular base pair, which were refined as two-position disorders (Figs. 2*b* and 2*c*).

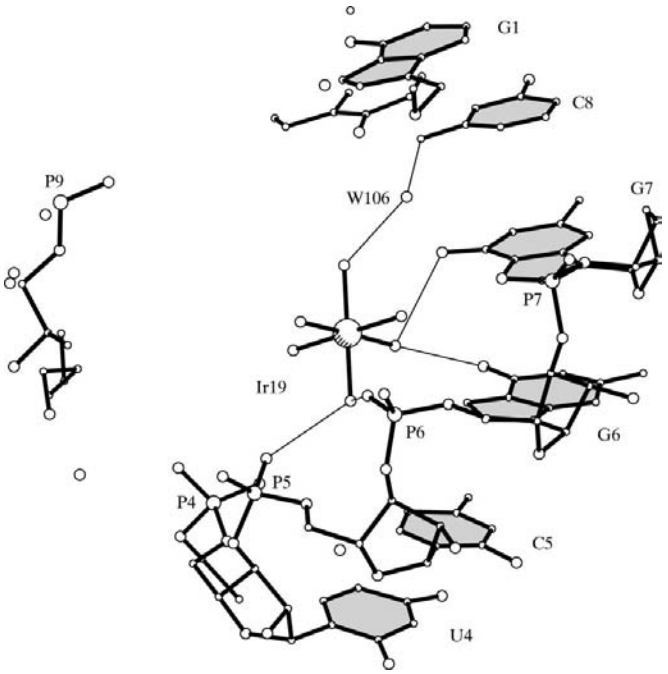

Figure 3
The network around the $M(NH_3)$ cation (site 1, M = Rh or I). The cation stands in a cavity with strong interactions bridging phosphate O atoms of C5 and G6, and also connect the corresponding guanine bases.

The central part of the duplex consists of four continuous non-Watson–Crick base pairs G–U and C–U. This region, though predicted to be less stable, is in fact well defined, with thermal factors notably lower than those of the adjacent Watson–Crick base pairs. The C–U or G–U base pairs, when including their structural water molecules in their binding scheme, mimic the geometry of standard A–U and G–C pairs. Thus, the C1—C1 vector angles and turn per base step parameters are very similar to those of a classic A-helix. As such, this central non-complementary core has a negligible effect on the deformation of the double helix compared with the 'standard' A duplex. In addition, the role of the two additional water molecules W#103 and W#113 in bridging the O2′ sugar atom to the facing base is of particular importance in the stability of the G–U mismatch in RNA compared with the G–T mismatch in DNA.

The four structures contain similar kinds of ordered binding sites for the hexammine cations, with clearly defined interactions between the oligonucleotide structure and the ammine groups of the cation.

The first site (Fig. 3) involves the phosphate O1P and O2P O atoms and the N7 and O6 atoms of consecutive guanine residues on the same strand.

The structure contains a hexammine cation-binding site of a second kind involving two adjacent C–U base pairs in the central region. The C–U base pairs are linked directly by one hydrogen bond C(N4)—H···(O4)U and indirectly by a hydrogen bond involving a bridging water molecule, U(N3)—H···OH2···(N3)C between the N3 atoms of the two pyrimidine rings. The free carbonyl groups of the two C–U pairs form strong hydrogen bonds to a hexammine cation-binding site on the minor-groove side of the duplex, with hydrogen-bond interactions between the carbonyl O atoms of cytosine and the ammine groups of the cation. This site is similar to the binding site for cobalt hexammine reported in the structure of d(CGCGCG), which has a Z conformation (Ho *et al.*, 1987).

The cation Mg^{2+} has considerable popular use in oligonucleotide crystallizations, partly because of its biological relevance and partly because of the high success rate in obtaining crystals. Mg^{2+} exists in solution mainly as $Mg(H_2O)_6^{2+}$, with six water O atoms arranged octahedrally about the Mg^{2+} ion at a nominal Mg—O distance of 2.0 Å. This situation is very similar to that which exists in hexammine cations, which have six covalently bound ammine groups arranged octahedrally about a central Co (Rh, Ru or Ir) atom. Under these conditions, one would expect a similar propensity to form bridges between negatively charged phosphates, leading to crystal formation.

The failure to identify Mg^{2+}-binding sites in most of the cases is consistent with general disorder of the cations in the thermally more mobile part of the water/cation network around the helices; usually, replacement of Mg^{2+} by an ion of higher atomic number is unlikely to produce useful derivatives for isomorphous replacement purposes. However, some success with the technique has been reported in either small (Cruse *et al.*, 1994) or large RNA systems (Ennifar *et al.*, 1999;

Wimberly *et al.*, 2000) with Os, Ru or Ir complexes as an alternative to the halogen method (incorporation of a iodocytosine or a bromouracil). One of the major drawbacks of the method is changes in unit-cell parameters, which is understandable as the ionic radius increases, thus introducing the classic non-isomorphism complication. These unfavourable changes arising from cation sizes were not observed in the four present structures, which all show identical unit-cell parameters, good isomorphism and ordered binding sites for all the cations, whatever the metal in use. As mentioned before, the non-isomorphism we observed in these crystal structures is only correlated with the presence of the fluorouracil extra DNA base.

In the initially solved structure, X = Br, M = Rh (PDB code 165d), two additional strong density peaks were tentatively interpreted as disordered secondary binding sites for hexammine cations with low occupancy factors. In the present high-resolution structures, from a reinspection of the electron density of the iridium derivative at 1.53 Å it was evident that these round-shaped peaks clearly correspond to single negatively charged atoms, as they lie in the vicinity of positively charged sites. They were refined as chloride ions, the counterion associated with the hexammine cations. These anionic sites are common to all structures.

3.1. The destabilizing effect of the fluorouracil

The double-stranded structure is not symmetric. In spite of equivalent character for the extra DNA halogenouracils at both ends, the cavities where they lie are different.

The d-^{X}U9 residue is located in a small pocket delimited by the minor groove of the symmetry-related stacked duplex. It is tightly bound by a number of polar contacts and hydrogen bonds, a situation that also prevails for all other halogenouracils (Fig. 2). At the other end, d-^{X}U18 stands in a larger pocket with more degrees of freedom. This last base plays an important role in the extension of the duplex. Owing to a balance between polar contacts, which forces the base out of the helix, and stacking, which pushes the base into the helical stack, the halogen (X) atom of the d-^{X}U18 residue thus modulates the length of the duplex. The halogen implies different electronic effects depending on its nature. A Br atom (and to a lesser extend a Cl atom) corresponds in size and hydrophobic effect to a methyl group, as observed in crystal structures of thymine-containing DNAs compared with their ^{Br}U analogues (Kennard *et al.*, 1986). In contrast, an F atom will have a strong electronic effect and plays a special role in biochemistry, not only in fluorouracil itself, which is widely used in anti-cancer therapy, but also in many other kind of inhibitors, whether peptidic (Li de la Sierra *et al.*, 1990; Parisi & Abeles, 1992; Ippolito & Christianson, 1992) or not (Palanki *et al.*, 2000). When switching from bromine or chlorine to fluorine, the hydrophobic effect reverts to a strong polar effect, inducing important interactions with the phosphate backbones of symmetry-related duplexes and ejection of the uracil from the stack.

References

Carter, R. J., Baeyens, K. J., Santalucia, J., Turner, D. H. & Holbrook, S. R. (1997). *Nucleic Acids Res.* **25**, 4117–4122.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cruse, W. B. T., Saludjian, P., Biala, E., Strazewski, P., Prangé, T. & Kennard, O. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 4160–4164.

Cruse, W. B. T., Saludjian, P., Leroux, Y., Léger, G., El Manouni, D. & Prangé, T. (1996). *J. Biol. Chem.* **271**, 15558–15566.

Davis, P. W., Thurmes, W. & Tinoco, I. Jr (1993). *Nucleic Acids Res.* **21**, 537–545.

Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C. & Dumas, P. (1999). *Structure*, **7**, 1439–1449.

Fourme, R., Dhez, P., Benoit, J. P., Kahn, R., Dubuisson, J. M., Besson, P. & Frouin, J. (1992). *Rev. Sci. Instrum.* **63**, 982–987.

Gao, Y. G., Liaw, Y. C., Robinson, H. & Wang, H. J. (1990). *Biochemistry*, **29**, 10307–10316.

Geissner, R. V., Quigley, G. J, Wang, H. J., van der Marel, A., van Boom, J. H. & Rich, A. (1985). *Biochemistry*, **24**, 237–240.

Hingerty, B. E., Brown, R. S. & Klug, A. (1982). *Biochim. Biophys. Acta*, **697**, 78–82.

Ho, P. S., Frederick, C. A., Saal, D., Wang, A. H. & Rich, A. (1987). *J. Biomol. Struct. Dyn.* **5**, 521–534.

Holbrook, S., Cheong, C., Tinoco, I. & Kim, S.-H. (1991). *Nature (London)*, **353**, 579–581.

Ippolito, J. A. & Christianson, D. W. (1992). *Int. J. Biol. Macromol.* **14**, 193–197.

Kennard, O., Cruse, W., Nachman, J., Prangé, T., Shakked, Z. & Rabinovitch, D. (1986). *J. Biomol. Struct. Dyn.* **3**, 623–647.

Leslie, A. G. W. (1994). *MOSFLM User Guide, MOSFLM Version* 5.20. MRC Laboratory of Molecular Biology, Cambridge.

Li de la Sierra, I., Papamichael, E., Sakarellos, C., Dimicoli, J. L. & Prangé, T. (1990). *J. Mol. Recogn.* **3**, 36–44.

Mirmira, S. R. & Tinoco, I. Jr (1996). *Biochemistry*, **35**, 7675–7683.

Navaza, J., Mauguen, Y., Saludjian, P., Prangé, T., Alzari, P. & Bentley, G. (1993). *Acta Cryst.* A**49**, C49.

Navaza, J., Pabepucci, E. H. & Martin, C. (1998). *Acta Cryst.* D**54**, 817–821.

Nunn, C. M. (1996). *J. Biomol. Struct. Dyn.* **14**, 49–56.

Palanki, M. S., Erdman, P. E., Gayo-Fung, L. M., Shevlin, G. I., Sullivan, R. W., Goldman, M. E., Ransone, L. J., Bennett, B. L., Manning, A. M. & Suto, M. J. (2000). *Med. Chem.* **43**, 3995–4004.

Parisi, M. F. & Abeles, R. H. (1992). *Biochemistry*, **31**, 9429–9435.

Sheldrick, G. (1997). *SHELX97. Program for Refinement of Crystal Structures.* University of Göttingen, Germany.

Timsit, Y. & Moras, D. (1994). *EMBO J.* **13**, 2737–2746.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M. Jr, Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. & Ramakrishnan, V. (2000). *Nature (London)*, **407**, 327–339.