Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

An unusual *syn* conformation of 5-formyluracil stabilized by supramolecular interactions

Gustavo Portalone* and Marcello Colapietro

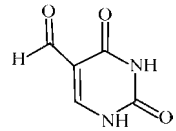
Chemistry Department, 'Sapienza' University of Rome, P.le A. Moro 5, I-00185 Rome, Italy
Correspondence e-mail: g.portalone@caspur.it

Received 30 July 2007
Accepted 17 September 2007
Online 13 October 2007

The asymmetric unit of the amino–oxo tautomer of 5-formyluracil (systematic name: 2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carbaldehyde), $C_5H_4N_2O_3$, comprises one planar amino–oxo tautomer, as every atom in the structure lies on a crystallographic mirror plane. At variance with all the previously reported small-molecule crystal structures containing the 5-formyluracil residue, the formyl substituent in the title compound exhibits an unusual *syn* conformation. The molecules are linked into planar sheets parallel to the *bc* plane by a combination of six N—H···O and C—H···O hydrogen bonds. Four of the hydrogen bonds are utilized to stabilize the formyl group in the *syn* conformation.

Comment

The discovery of double helix, as Watson and Crick realised (Watson & Crick, 1953), immediately provided fundamental new insights into the nature of genetic events. Our more recent knowledge of both the detail and the variety of DNA structures themselves, and the manner in which they are recognized by regulatory proteins, mutational compounds and drugs, is starting to pave the way to more profound levels of understanding of the processes of gene regulation, mutation/carcinogenesis and drug action at the molecular level. However, despite improvements in the average resolution of crystal structures, there remains a need to clarify structural details for the better understanding of structure–function and structure–stability relationships. The underlying relationships between DNA sequences, structure and flexibility, even though only partially understood, are of great interest for understanding the interaction of both small molecules and proteins with DNA, as well as for understanding the function of the resulting complexes. In particular, 5-formyluracil is known to miscode with relatively high frequency, generating primarily transition mutations (Bjelland *et al.*, 1995, 2001). Consequently, oligonucleotides containing the 5-formyluracil residue, a major oxidation product of thymine, have been investigated extensively by a variety of experimental methods in the last decade. The strong miscoding potential of 5-formyluracil has been attributed to the strong electron-withdrawing 5-formyl substituent, which substantially increases the acidity of the N-bound H atoms (Sowers *et al.*, 1987; Privat & Sowers, 1996; Jang *et al.*, 2001). In addition to altering the electronic properties of the pyrimidine bases, the oxidation of the thymine methyl group to the 5-formyl substituent would be expected to interfere with sequence-specific DNA–protein interactions (Rogstad *et al.*, 2004), and to form potentially lethal covalent crosslinks between the 5-formyl group and the amino groups of DNA-binding proteins, resulting in locking the binding (Armstrong, Sternbach & Eckstein, 1976; Mee & Adelstein, 1981; Ono *et al.*, 1994; Sugiyama *et al.*, 2001; Rogstad *et al.*, 2002).



(I)

Like other related DNA/RNA pyrimidine bases (Saenger, 1984), 5-formyluracil [5formur, (I)] exhibits a tautomeric keto–enol equilibrium. The tautomeric equilibrium of nucleobases is important, owing to the strong relations between the potential coding properties and the tautomeric form. A few years ago, the crystal structure of a Dickerson–Drew-type dodecamer with the sequence d(CGCGAAT*X*CGCG), containing 2′-deoxy-5-formyluridine at *X*, was determined (Tsunoda *et al.*, 2002). In this study, the formyl group of one of the two keto tautomers of the 5-formyluracil residues adopts a *syn* conformation, and the second is disordered between the *syn* and *anti* conformations with almost equal occupancies. Both tautomers form two hydrogen bonds to opposite adenine residues, yielding canonical WC base pairs in the same way as thymine. At variance with the previous results, it has been proposed that the enol form of 5formur might interact with thymine forming a WC mispair (Ånensen *et al.*, 2001). Theoretical studies at different levels of theory have been reported concerning 5-formyluracil in the gas phase and in polar and nonpolar solvent fields (Cysewski *et al.*, 1998; Jang *et al.*, 2001). It has been shown that the diketo form, with the formyl group oriented so as to preserve the maximum distance between adjacent O2 and O7 atoms (*anti* conformation of the formyl moiety), is preferred by 4.5 kcal mol^{-1} (1 kcal mol^{-1} = 4.184 kJ mol^{-1}) in the gas phase but has a 4.6 kcal mol^{-1} lower solvation energy in water, leading to similar energies for both conformers in solution (the difference being less than 0.1 kcal mol^{-1}). Therefore, a particular interest in the crystal structure of 5-formyluracil is to see which tautomeric form is present and which conformation is adopted by the formyl moiety.

In this context, as part of our continuing study of crystal adducts of DNA/RNA pyrimidine bases coupled with amino derivatives of aromatic N-heterocycles *via* multiple hydrogen bonds to mimic the base pairing of nucleic acids (Portalone *et al.*, 1999, 2002; Brunetti *et al.*, 2000, 2002; Portalone & Colapietro, 2004*a*,*b*, 2006, 2007*a*,*b*,*c*,*d*; Portalone, 2007) and in

  DOI: 10.1107/S0108270107045659

view of the great importance of the modified nucleobase 5-formyluracil, we have been attracted by the crystal structure of 5formur, surprisingly not yet reported.

In the crystal structure of (I), the asymmetric unit comprises one amino–oxo tautomer, with every atom constrained to lie on a crystallographic mirror plane (Fig. 1). The principal point of interest in the molecular structure of 5formur is the unusual *syn* conformation adopted by the 5-formyl substituent. A search of the Cambridge Structural Database (Version 5.28; Allen, 2002) for crystal structures containing the 5-formyluracil and 5-formyl-1-thiouracil residues yields only four structures (Armstrong, Dattagupta *et al.*, 1976; Garcia-Megias *et al.*, 1989; Hernández *et al.*, 1997; Kittaka *et al.*, 2004). For all these compounds, the 5-formyl group exists in an *anti* conformation, in agreement with the previously discussed theoretical results for the free molecule. The 'unexpected' *syn* conformation of the formyl group in 5formur causes almost identical but opposite variations of the exocyclic bond angles C4—C5—C7 and C6—C5—C7 [4.6–4.8 (3)°] with respect to the corresponding bond angles in thymine (Portalone *et al.*, 1999), but leaves the bond angle at the *ipso* C atom, C4—C5—C6, unchanged (Table 1). No dependence of the *ipso* angle on the conformation adopted by the formyl group has been found. The lack of an appreciable angular distortion at the *ipso* angle, a parameter that is particularly sensitive to the σ-inductive effect of the substituent, is rather surprising. This angle is 120.5 (4)° from a study of the 1:1 *syn–anti* conformers of terephthalaldehyde in the gas phase (Bock *et al.*, 1987), and this value is intermediate between those obtained by electron diffraction for *p*-xylene [117.1 (3)°; Domenicano *et al.*, 1979] and *p*-difluorobenzene [123.5 (1)°; Domenicano *et al.*, 1982], in accordance with the values of Taft's inductive parameter, σ_I, which are −0.08, 0.25 and 0.52 for the CH_3, CHO and F substituents, respectively (Taft *et al.*, 1963).

A further comparison of the molecular geometry of 5formur with that reported for uracil and thymine (Stewart & Jensen, 1967; Portalone *et al.*, 1999) should clarify the influence exerted by the strong electron-withdrawing 5-formyl substituent on the overall molecular electronic structure. However, from this comparison no clear indications point to the importance of charge-separated quininoid forms as significant contributors. Consequently, the reported small geometric distortions (see Δ values in Table 1) have been attributed to different hydrogen-bonding configurations. On the other hand, it has recently been debated whether generally accepted resonance forms can be applied to explain the structure of pyrimidinic nucleobases (Gonzáles Moa & Mosquera, 2003, 2005). These authors, from a topological analysis of charge densities carried out with the atoms in molecule (AIM) theory (Bader, 1990), came to the conclusion that the resonance model does not correctly describe the charge distribution in the neutral (and protonated) forms of uracil derivatives.

The hydrogen-bonding arrangement in (I), shown in Fig. 2, is slightly unusual, in that all hydrogen-bond donors and acceptors are involved in the hydrogen-bonded sheet structure (Table 2). It is far more commonly observed in the packing of 5-substituted uracils, where the C5-substituents are limited to those belonging to the second row of the periodic table and unable to form internal hydrogen-bonding interactions, that one carbonyl group is partially unsaturated (Jeffrey & Saenger, 1991). In 5formur, a total of six two- and three-centre N—H···O and C—H···O hydrogen bonds, four involving the formyl substituent, delineate patterns in which hydrogen-bonded noncentrosymmetric synthons are the prominent features. An asymmetric bifurcated N—H···O interaction, $R_1^2(6)$ (Etter *et al.*, 1990; Bernstein *et al.*, 1995; Motherwell *et al.*, 1999), connects molecules related by a translational operation and induces the formation of infinite polar chains along the [010] direction. These infinite chains are then cross-linked by one N—H···O, one C—H···O and one bifurcated C—H···O intermolecular hydrogen bond, forming a sheet-like structure *via* four adjoining hydrogen-bonded rings [$R_3^2(9)$, $R_2^2(7)$, $R_1^2(6)$ and $R_2^1(6)$]. The stacking of these planar layers, spaced in the *a*-axis direction by ∼4.6 Å, produces relatively little overlap of the molecules. In one of the three N—H···O interactions, where atom N1 acts as hydrogen-bond

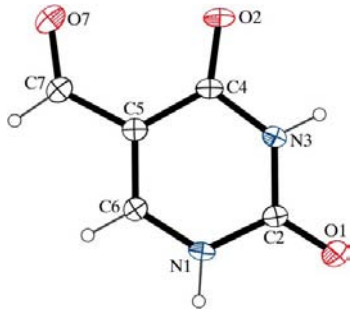


Figure 1
The crystallographic asymmetric unit in (I), showing the atom-labelling scheme. Displacement ellipsoids are drawn at the 50% probability level and H atoms are shown as small spheres of arbitrary radii.

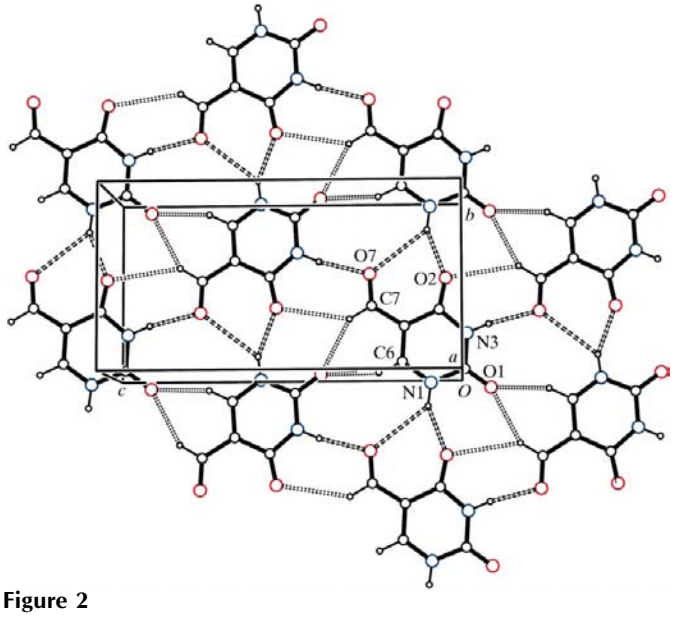


Figure 2
A crystal packing diagram for (I), viewed down *a*. All atoms are shown as small spheres of arbitrary radii. For the sake of clarity, N—H···O hydrogen bonds are indicated by dashed lines and C—H···O hydrogen bonds by dotted lines.

donor *via* atom H1, and one of the three C—H···O interactions, where atom C7 acts as hydrogen-bond donor *via* atom H7, there is some uncertainty as to whether they are hydrogen bonds or not. However, as is very frequently found for bifurcated hydrogen bonds, the sum of the inter-bond angles at the H atom is close to 360° and the H···O distance can be greater than the van der Waals separation (Jeffrey & Saenger, 1991; Desiraju & Steiner, 1999; Steiner, 2002).

As previously mentioned, modifications of DNA bases, which can significantly alter the ionization constant of H atoms involved in hydrogen-bond formation, are correlated with increased base mispair formation. The interpretation of the experimental evidence for the changes in pK_a with substitution at the 5-position in uracil has been proposed by spectroscopic methods in solution (Privat & Sowers, 1996), and theoretical investigations in polar solvents, based on first principles in quantum mechanics (density functional theory at the B3LYP level; Becke, 1993), have been used to take into account the presence in the heterocyclic ring of two sites (N1 and N3) for ionization (Jang *et al.*, 2001). From these studies, an elegant linear relationship was observed between the electronic inductive property of the 5-substituent and the experimental average pK_a value, and it was possible to resolve which cases prefer ionization at N1 or N3. In particular, it was proposed that, in addition to stabilizing the anionic form by inductive effects, the formyl and nitro substituents at C5 can alternatively stabilize the resonance ionization of the N1 H atom. This interpretation seems not to apply to 5-substituted uracils in the solid state: intermolecular hydrogen bonds should be reinforced by the existence of resonance forms which delocalize negative charge extensively (Jeffrey, 1997), and their importance should be related in turn to the acidity of the N atoms. However, no clear correlation has been found between the pK_a values of atoms N1 and N3 and the hydrogen-bond geometries (Table 3). Recently, topological factors have been invoked to explain the formation of the complex [Zn(uracilate-κN^1)(uracilate-κN^3)$(NH_3)_2$], where the unprecedented combination of both N1 and N3 uracil tautomers with the same metal atom demonstrates the overriding importance of noncovalent interactions in tautomer selection during crystallization (Escorihuela *et al.*, 2004).

It has been clearly demonstrated that the *anti* conformer of terephthalaldehyde in the gas phase is more stable than the *syn* conformer by a mere 0.21 kcal mol^{-1}, the *syn–anti* ratio being 1:1 within experimental error (Bock *et al.*, 1987). A similar situation has been predicted for 5-formyluracil in solution, and again the *syn–anti* ratio was 1:1 (Jang *et al.*, 2001). From these results it seems possible to foresee the existence of polymorphic crystals of 5-formyluracil in the *anti* conformation, and to design cocrystals or solvate structures where the existence of *syn*/*anti* conformational isomers of 5-formyluracil is pre-organized by different supramolecular aggregations (Hofmann *et al.*, 2004). This issue would be interesting to pursue in the future, with the help of computational analysis of hypothetical crystal structures of 5-formyluracil at lower lattice energies, analogous to those performed for uracil (Price & Wibley, 1997) and 5-fluorouracil (Hulme *et al.*, 2005).

Experimental

Very small crystals of the title compound (purchased from Sigma Aldrich at 98% purity) were obtained without further purification from a solution in water by slow evaporation of the solvent.

Crystal data

$C_5H_4N_2O_3$
M_r = 140.10
Orthorhombic, $Cmc2_1$
a = 6.2840 (5) Å
b = 6.7050 (6) Å
c = 12.8302 (9) Å
V = 540.59 (8) Å^3
Z = 4
Mo $K\alpha$ radiation
μ = 0.15 mm^{-1}
T = 298 (2) K
0.20 × 0.10 × 0.10 mm

Data collection

Huber CS single-crystal diffractometer
710 measured reflections
413 independent reflections
398 reflections with $I > 2\sigma(I)$
R_{int} = 0.046
3 standard reflections every 97 reflections
intensity decay: 2%

Refinement

$R[F^2 > 2\sigma(F^2)]$ = 0.043
$wR(F^2)$ = 0.112
S = 1.08
413 reflections
61 parameters
1 restraint
H-atom parameters constrained
$\Delta\rho_{max}$ = 0.29 e Å^{-3}
$\Delta\rho_{min}$ = −0.26 e Å^{-3}

Table 1
Selected geometric parameters (Å, °) for thymine, uracil and 5-formyluracil.

Δ_1 is the difference (Å) between corresponding bond distances in thymine and uracil, and Δ_2 is the difference (Å) between corresponding bond distances in 5-formyluracil and uracil.

	Thymine[a]	Uracil[b]	5Formur[c]	Δ_1	Δ_2
O1—C2	1.244 (4)	1.216 (2)	1.207 (4)	0.028 (4)	−0.009 (4)
O2—C4	1.225 (4)	1.242 (2)	1.220 (3)	−0.017 (4)	−0.022 (3)
N1—C2	1.358 (4)	1.373 (2)	1.386 (4)	−0.015 (4)	0.013 (4)
N1—C6	1.384 (5)	1.363 (2)	1.343 (4)	0.021 (5)	−0.020 (4)
N3—C2	1.361 (4)	1.378 (2)	1.358 (4)	−0.017 (4)	−0.020 (4)
N3—C4	1.401 (5)	1.374 (2)	1.384 (4)	0.027 (5)	0.010 (4)
C4—C5	1.453 (4)	1.432 (2)	1.458 (4)	0.021 (4)	0.026 (4)
C5—C6	1.343 (4)	1.346 (2)	1.360 (3)	−0.003 (4)	0.014 (3)
C5—C7	1.502 (6)		1.456 (5)		
C7—O7			1.215 (4)		
C4—C5—C6	118.4 (3)	118.9 (2)	118.6 (3)		
C4—C5—C7	118.3 (3)		122.9 (3)		
C6—C5—C7	123.3 (3)		118.5 (3)		

References: (*a*) Portalone *et al.* (1999); (*b*) Stewart & Jensen (1967); (*c*) this work.

Table 2
Hydrogen-bond and short-contact geometry (Å, °) for (I).

D—H···A	D—H	H···A	D···A	D—H···A
N1—H1···O2^{i}	0.89	1.98	2.794 (3)	152
N1—H1···O7^{i}	0.89	2.75	3.442 (4)	135
N3—H3···O7^{ii}	0.89	1.99	2.868 (4)	171
C6—H6···O1^{iii}	0.98	2.18	3.108 (5)	157
C7—H7···O1^{iii}	0.98	2.36	3.231 (5)	148
C7—H7···O2^{iv}	0.98	2.80	3.613 (5)	140

Symmetry codes: (i) $x, y-1, z$; (ii) $-x, -y+1, z-\frac{1}{2}$; (iii) $-x, -y, z+\frac{1}{2}$; (iv) $-x, -y+1, z+\frac{1}{2}$.

Table 3
Comparison of hydrogen-bonding geometry (Å, °) versus pK_a for thymine[a], uracil[b], 5-fluorouracil[c], 5-formyluracil[d] and 5-nitrouracil[e,f].

The pK_a values reported for 5-formyluracil refer to the *syn* conformer.

C5 Substituent	pK_a (experimental)[g]	pK_a(N1) (calculated)[h]	N1···O	N1—H1···O	pK_a(N3) (calculated)[h]	N3···O	N3—H3···O
CH_3[a]	9.75	11.23	2.827 (3)	178 (3)	10.04	2.833 (3)	175 (3)
H[b]	9.42	10.47	2.861 (2)	171 (1)	9.34	2.862 (3)	175 (1)
F (Molecule *A*)[c]	7.93	9.05	2.858 (2)	175 (2)	7.26	2.838 (2)	172 (2)
F (Molecule *B*)[c]	7.93	9.05	2.823 (2)	170 (2)	7.26	2.831 (2)	173 (2)
F (Molecule *C*)[c]	7.93	9.05	2.787 (2)	175 (2)	7.26	2.818 (2)	174 (2)
F (Molecule *D*)[c]	7.93	9.05	2.815 (2)	176 (2)	7.26	2.837 (2)	171 (2)
CHO[d]	6.84	6.95	2.794 (3)	152 (2)	7.28	2.869 (4)	171 (2)
NO_2, $P2_1/n$[e]	5.30	5.66	2.873 (2)	167 (2)	6.91	2.890 (2)	170 (2)
NO_2, $P2_12_12_1$[f]	5.30	5.66	2.790 (1)	166 (2)	6.91	2.793 (1)	156 (2)
NO_2, *Pbca*[f]	5.30	5.66	2.860 (1)	170 (2)	6.91	2.820 (1)	175 (2)

References: (*a*) Portalone *et al.* (1999); (*b*) Stewart & Jensen (1967); (*c*) Hulme *et al.* (2005), $P\overline{1}$ with four molecules in the asymmetric unit; (*d*) this work; (*e*) Kennedy *et al.* (1998); (*f*) Srinivasa Gopalan *et al.* (2000); (*g*) Privat & Sowers (1996); (*h*) Jang *et al.* (2001).

Diffraction from the very small crystals was weak; nevertheless, these data gave good structural results, albeit with a lower data/parameter ratio than usual. All H atoms were revealed by a difference synthesis calculated after the first cycles of isotropic refinement. All H atoms were positioned with idealized geometry and refined using a riding model, with distances C—H = 0.93 Å and N—H = 0.86 Å, and with U_{iso}(H) = 1.2U_{eq}(C,N). In the absence of significant anomalous scattering in this light-atom study, Friedel pairs were merged.

Data collection: *XCS* (Colapietro *et al.*, 1992); cell refinement: *XCS*; data reduction: *XCS*; program(s) used to solve structure: *SIR97* (Altomare *et al.*, 1999); program(s) used to refine structure: *SHELXL97* (Sheldrick, 1997); molecular graphics: *ORTEP-3* (Farrugia, 1997); software used to prepare material for publication: *SHELXL97*.

The authors thank MIUR (Rome) for financial support in 2006 of the project 'X-ray Diffractometry and Spectrometry'.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: HJ3051). Services for accessing these data are described at the back of the journal.

References

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Altomare, A., Burla, M. C., Camalli, M., Cascarano, G. L., Giacovazzo, C., Guagliardi, A., Moliterni, A. G. G., Polidori, G. & Spagna, R. (1999). *J. Appl. Cryst.* **32**, 115–119.
Ånensen, H., Provan, F., Lian, A. T., Reinertsen, S.-H. H. S., Ueno, Y., Matsuda, A., Seeberg, E. & Bjelland, S. (2001). *Mutat. Res.* **476**, 99–107.
Armstrong, V. W., Dattagupta, J. K., Eckstein, F. & Saenger, W. (1976). *Nucleic Acids Res.* **3**, 1791–1810.
Armstrong, V. W., Sternbach, H. & Eckstein, F. (1976). *Biochemistry*, **15**, 2086–2091.
Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory*. New York: Oxford University Press.
Becke, A. D. (1993). *J. Phys. Chem.* **98**, 5649–5652.
Bernstein, J., Davis, R. E., Shimoni, L. & Chang, N.-L. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555–1573.
Bjelland, S., Ånensen, H., Knaevelsrud, I. & Seeberg, E. (2001). *Mutat. Res.* **486**, 147–154.
Bjelland, S., Eide, L., Time, R. W., Stote, R., Eftedal, I., Volden, G. & Seeberg, E. (1995). *Biochemistry*, **34**, 14758–14764.
Bock, C. W., Domenicano, A., George, P., Hargittai, I., Portalone, G. & Schultz, Gy. (1987). *J. Phys. Chem.* **91**, 6120–6127.
Brunetti, B., Piacente, V. & Portalone, G. (2000). *J. Chem. Eng. Data*, **45**, 242–246.
Brunetti, B., Piacente, V. & Portalone, G. (2002). *J. Chem. Eng. Data*, **47**, 17–19.
Colapietro, M., Cappuccio, G., Marciante, C., Pifferi, A., Spagna, R. & Helliwell, J. R. (1992). *J. Appl. Cryst.* **25**, 192–194.
Cysewski, P., Jeziorek, D. & Oliński, R. (1998). *J. Chem. Soc. Faraday Trans.* **94**, 1813–1821.
Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology*, pp. 19–21. Oxford University Press.
Domenicano, A., Kolonits, M. & Hargittai, I. (1982). *J. Mol. Struct.* **78**, 97–111.
Domenicano, A., Schultz, Gy., Kolonits, M. & Hargittai, I. (1979). *J. Mol. Struct.* **53**, 197–209.
Escorihuela, I., Falvello, L. R., Tomás, M. & Urriolabeitia, E. P. (2004). *Cryst. Growth Des.* **4**, 655–657.
Etter, M. C., MacDonald, J. C. & Bernstein, J. (1990). *Acta Cryst.* B**46**, 256–262.
Farrugia, L. J. (1997). *J. Appl. Cryst.* **30**, 565.
Garcia-Megias, L. E., Colacio-Rodriguez, E., Salas-Peregrin, J. M., Simard, M. & Beauchamp, A. L. (1989). *Inorg. Chim. Acta*, **158**, 149–153.
Gonzáles Moa, M. J. & Mosquera, R. A. (2003). *J. Phys. Chem. A*, **107**, 5361–5367.
Gonzáles Moa, M. J. & Mosquera, R. A. (2005). *J. Phys. Chem. A*, **109**, 3682–3686.
Hernández, R. P., Morales, A. D., Mascarenhas, Y., Carretero, M. M., Pulido, S. J. & Ureña, F. H. (1997). *Acta Cryst.* C**53**, IUC9700012.
Hofmann, D. W. M., Kuleshova, L. N. & Antipin, M. Yu. (2004). *Cryst. Growth Des.* **4**, 1395–1402.
Hulme, A. T., Price, S. L. & Tocher, D. A. (2005). *J. Am. Chem. Soc.* **127**, 1116–1117.
Jang, Y. H., Sowers, L. C., Çağin, T. & Goddard, W. A. III (2001). *J. Phys. Chem. A*, **105**, 274–280.
Jeffrey, G. A. (1997). *An Introduction to Hydrogen Bonding*, pp. 98–100. New York: Oxford University Press.
Jeffrey, G. A. & Saenger, W. (1991). *Hydrogen Bonding in Biological Structures*, pp. 247–268. Berlin/Heidelberg: Springer-Verlag.
Kennedy, A. R., Okoth, M. O., Sheen, D. B., Sherwood, J. N. & Vrcelj, R. M. (1998). *Acta Cryst.* C**54**, 547–550.
Kittaka, A., Horii, C., Tanaka, H., Miyasaka, T., Nakamura, K. T., Kuroda, R. & Sugiyama, T. (2004). *Heterocycles*, **64**, 367–382.
Mee, L. K. & Adelstein, S. J. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 2194–2198.
Motherwell, W. D. S., Shields, G. P. & Allen, F. H. (1999). *Acta Cryst.* B**55**, 1044–1056.
Ono, A., Okamoto, T., Inada, M., Nara, H. & Matsuda, A. (1994). *Chem. Pharm. Bull.* **42**, 2231–2237.
Portalone, G. (2007). *Acta Cryst.* E**63**, o3232.
Portalone, G., Ballirano, P. & Maras, A. (2002). *J. Mol. Struct.* **608**, 35–39.
Portalone, G., Bencivenni, L., Colapietro, M., Pieretti, A. & Ramondo, F. (1999). *Acta Chem. Scand.* **53**, 57–68.
Portalone, G. & Colapietro, M. (2004*a*). *Acta Cryst.* E**60**, o1165–o1166.
Portalone, G. & Colapietro, M. (2004*b*). *J. Chem. Crystallogr.* **34**, 609–612.
Portalone, G. & Colapietro, M. (2006). *Acta Cryst.* E**62**, o1049–o1051.

Portalone, G. & Colapietro, M. (2007*a*). *Acta Cryst.* C**63**, o181–o184.
Portalone, G. & Colapietro, M. (2007*b*). *Acta Cryst.* E**63**, o1869–o1871.
Portalone, G. & Colapietro, M. (2007*c*). *J. Chem. Crystallogr.* **37**, 141–145.
Portalone, G. & Colapietro, M. (2007*d*). *Acta Cryst.* C**63**, o423–o425.
Price, S. L. & Wibley, K. S. (1997). *J. Phys. Chem. A*, **101**, 2198–2206.
Privat, E. J. & Sowers, L. C. (1996). *Mutat. Res.* **354**, 151–156.
Rogstad, D. K., Heo, J., Vaidehi, N., Goddard, W. A. III, Burdzy, A. & Sowers, L. C. (2004). *Biochemistry*, **43**, 5688–5697.
Rogstad, D. K., Liu, P., Burdzy, A., Lin, S. S. & Sowers, L. C. (2002). *Biochemistry*, **41**, 8093–8102.
Saenger, W. (1984). *Principles of Nucleic Acid Structure*, pp. 110–115. New York: Springer.
Sheldrick, G. M. (1997). *SHELXL97*. University of Göttingen, Germany.
Sowers, L. C., Ramsay Shaw, B., Veigl, M. L. & Sedwick, W. D. (1987). *Mutat. Res.* **177**, 201–218.
Srinivasa Gopalan, R., Kulkarni, G. U. & Rao, C. N. R. (2000). *ChemPhysChem*, **1**, 127–135.
Steiner, T. (2002). *Angew. Chem. Int. Ed.* **41**, 48–76.
Stewart, R. F. & Jensen, L. H. (1967). *Acta Cryst.* **23**, 1102–1105.
Sugiyama, T., Kittaka, A., Takayama, H., Tomioka, M., Ida, Y. & Kuroda, R. (2001). *Nucleosides Nucleotides Nucleic Acids*, **20**, 1079–1083.
Taft, R. W., Price, E., Fox, I. R., Lewis, I. C., Andersen, K. K. & Davis, G. T. (1963). *J. Am. Chem. Soc.* **85**, 709–724.
Tsunoda, M., Kondo, J., Karino, N., Ueno, Y., Matsuda, A. & Takenaka, A. (2002). *Biophys. Chem.* **95**, 227–233.
Watson, J. D. & Crick, F. H. C. (1953). *Nature*, **171**, 737–738.