Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

The 1:1 complex of cytosine and 5-fluorouracil monohydrate revisited

Gustavo Portalone* and Marcello Colapietro

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

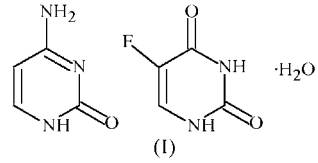
Received 23 April 2007
Accepted 31 May 2007
Online 23 June 2007

The monohydrated molecular adduct cytosine–5-fluorouracil–water (1/1/1) (denoted CytFur) [systematic name: 4-amino-pyrimidin-2(1*H*)-one–5-fluoropyrimidine-2,4(1*H*,3*H*)-dione–water (1/1/1)], $C_4H_5N_3O{\cdot}C_4H_3FN_2O_2{\cdot}H_2O$, was determined some 40 years ago [Voet & Rich (1969). *J. Am. Chem. Soc.* **91**, 3069–3075] and is widely cited as the first example of an intermolecular complex between two pyrimidinic nucleobases. In view of the importance of these base associations, CytFur has been reinvestigated with modern laboratory equipment to higher precision and with the location and free refinement of the H atoms. The new experiment reaffirms the results of the original and clarifies the tautomeric form exhibited by the compounds. The asymmetric unit comprises a hydrogen-bonded adduct of the canonical amino–oxo tautomers in an exact 1:1 ratio and a water molecule of crystallization. This cyclic dimer forms a layered structure approximately parallel to the *bc* plane by joining through hydrogen bonds other such cyclic dimers. Disordered water molecules run through tunnels formed by surrounding molecular adducts along the *a* axis.

Comment

In the past two decades, much attention has been paid to non-Watson–Crick base associations (mismatch) to further the understanding of their influence on the structure of duplex DNA and RNA (Hunter & Brown, 1999, and references therein). Among the eight nonstandard base pairings in RNA, the three pyrimidine–pyrimidine mismatches (C·C, U·U and C·U) have proved difficult to characterize, but there are a few examples of C·U and U·U base-pair incorporation in duplex RNA (Holbrook *et al.*, 1991). In particular, in the C·U mispair, there is a single N—H···O hydrogen bond involving the cytosine N6 and uracil O2 atoms, and a bridging water molecule links the two N3 groups. In noncanonical base pairing, these conventional hydrogen bonds can be flanked by weaker interbase C—H···O hydrogen bonds (Desiraju & Steiner, 1999), and they occur in a number of interactions involving modified bases and in triplexes. For example, in the crystal structure of the RNA hexamer UUCGCG, a mismatch U·U pair linked by an N—H···O and a C—H···O hydrogen bond was observed, the so-called 'Calcutta pair' (Wahl *et al.*, 1996). The theory of mispair association relies on the formation of rare tautomer forms of the bases (Strazewski & Tamm, 1990). However, there is no direct experimental evidence for these base pairs, owing to crystallographic resolution problems (Hunter & Brown, 1999).

The structure of the title compound (CytFur), (I), has been known for a long time as the first example of an intermolecular complex between two pyrimidinic nucleobases, *viz.* cytosine and the chemotherapeutic agent 5-fluorouracil, and the latter is well known to incorporate into DNA *in vivo* (Pinedo & Peters, 1988). CytFur was determined as a monohydrate some 40 years ago (Voet & Rich, 1969), with remarkable accuracy for that time. 1730 unique reflections were collected at ambient temperature by photographic techniques using Cu $K\alpha$ radiation and their intensities were estimated visually. Patterson methods were employed to solve the crystal structure, but only non-H atoms were positioned and refined. The final refinement, carried out on a fairly small data set (1189 observed reflections of non-zero weight), led to $R = 0.162$ with a data-to-parameter ratio of 6.9, and standard deviations of 0.01 Å in bond lengths and 1° in bond angles, but with a number of problems. Specifically, in the refinement, the matrices of the anisotropic displacement parameters for 5-fluorouracil atoms C4 and C5 and cytosine atom N4 were not positive definite quantities. This situation, together with the large shifts in the displacement parameters for these atoms and the high value of the residual after the final refinement, was attributed by the authors to the disorder of the water molecule [the sum of the unconstrained refined occupancy factors for the two disorder components was only close to unity, *viz.* 0.76 (11) and 0.26 (8)].



(I)

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 (Portalone *et al.*, 1999, 2002; Brunetti *et al.*, 2000, 2002; Portalone & Colapietro, 2004*a*,*b*, 2006, 2007*a*,*b*,*c*), we report here a reinvestigation of (I), with a full description of the molecular and supramolecular structures. Our results reaffirm Voet's work, providing also the locations of the H atoms, a lower R value, a data-to-parameter ratio of 14.2 and a significant increase in the precision of the geometric parameters.

The asymmetric unit of (I) consists of 1:1 hydrogen-bonded coplanar canonical amino–oxo tautomers of the two nucleobases and a water molecule (Fig. 1). The water molecule is distributed over two sites with ∼50:50 occupancy factors. A comparison of the geometric parameters of the molecular components of CytFur with those reported for cytosine monohydrate (Cytosm; McClure & Craven, 1973) and for 5-fluorouracil (5Furac; Hulme *et al.*, 2005) shows that the

 DOI: 10.1107/S0108270107026649

corresponding bond lengths and angles are equal within experimental error (Table 1), but with minor exceptions that can be attributed to different hydrogen-bonding configurations. For instance, in the crystal structure of CytFur, the hydrogen-bonding scheme involves all H-atom donor/acceptor sites of the pyrimidinic bases, at variance with that observed in 5-fluorouracil, where atom O1 remains partially unsaturated. As a consequence, the C2—O1 bond length in CytFur is slightly longer than the corresponding bond in 5Furac by 0.008 (1) Å.

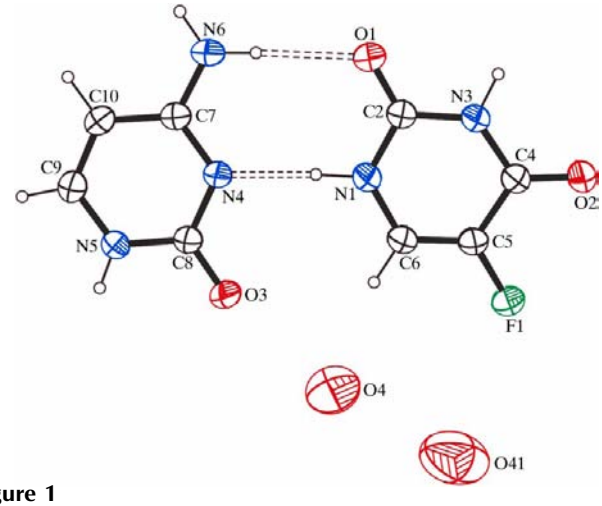


Figure 1
The crystallographic asymmetric unit in CytFur, showing the atom-labelling scheme and hydrogen bonding (dashed lines). Displacement ellipsoids are drawn at the 50% probability level and H atoms are shown as small spheres of arbitrary radii.

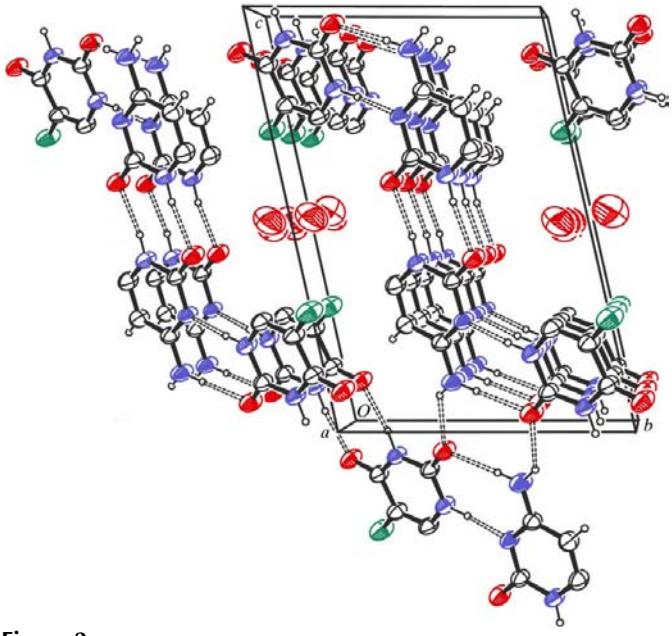


Figure 2
The crystal packing of CytFur, viewed down *a*. Displacement ellipsoids are drawn at the 50% probability level. For the sake of clarity, only H atoms involved in hydrogen bonding are shown (small spheres of arbitrary radii). Hydrogen bonding is indicated by dashed lines.

As previously mentioned, CytFur crystallizes as a monohydrated 1:1 molecular adduct. The supramolecular structure of (I) is dominated by two main motifs based on four structurally significant hydrogen bonds, one each of types O—H···O, N—H···O, N—H···N and C—H···O (Table 2). Firstly, an $R_2^2(8)$ ring (Etter *et al.*, 1990; Bernstein *et al.*, 1995; Motherwell *et al.*, 1999) forms from N—H···O and N—H···N double intermolecular hydrogen bonds between coplanar base pairs. Propagation and inversion of the base pair through $R_2^2(8)$ N—H···O double intermolecular hydrogen bonds generate infinite chains of centrosymmetric rings running approximately parallel to the [011] direction (Fig. 2). In other words, this hydrogen-bonding scheme corresponds to an alternating double repetition of cytosine and 5-fluorouracil residues. These infinite chains are then crosslinked by symmetry-related N—H···O intermolecular hydrogen bonds between two sets of cytosine and 5-fluorouracil molecules in a tetrameric arrangement and form the second major motif, also shown in Fig. 2, namely an $R_4^4(8)$ ring. The hydrogen bonds so far discussed build a two-dimensional array, which is modestly reinforced by the C10—H10···O2^{vi} [symmetry code: (vi) $x-1$, $y-1$, z] intermolecular interaction between adjacent cytosine and 5-fluorouracil molecules. Disordered water molecules run through channels formed by surrounding centrosymmetric molecular adducts along the *a* axis. As their H atoms could not be located, it was impossible to assess the involvement of water molecules in hydrogen bonding. Nevertheless, the distance between cytosine atom O3 and the positions of the partial water atom O4 [2.902 (3) Å], as well as the good agreement between the C8—O3 bond lengths of Cytfur and Cytosm (in Cytosm, atom O3 forms three hydrogen bonds, and two of them involve two water molecules) could suggest participation of water molecules in a weak hydrogen bond.

Experimental

Cytosine and 5-fluorouracil were purchased from Sigma–Aldrich (99% purity) and used as obtained. Crystals of CytFur were grown from an aqueous solution (0.1 mmol of each compound in *ca* 8 ml) by slow evaporation of the solvent.

Crystal data

$C_4H_5N_3O{\cdot}C_4H_3FN_2O_2{\cdot}H_2O$
M_r = 259.21
Triclinic, $P\overline{1}$
a = 4.2629 (1) Å
b = 9.5372 (5) Å
c = 13.9639 (9) Å
α = 102.562 (5)°
β = 91.055 (7)°
γ = 101.331 (7)°
V = 542.22 (5) Å^3
Z = 2
Mo $K\alpha$ radiation
μ = 0.14 mm^{-1}
T = 298 (2) K
0.20 × 0.20 × 0.15 mm

Data collection

Huber CS four-circle diffractometer
4833 measured reflections
4399 independent reflections
2901 reflections with $I > 2\sigma(I)$
R_{int} = 0.020
3 standard reflections
every 97 reflections
intensity decay: 1%

Refinement

$R[F^2 > 2\sigma(F^2)]$ = 0.047
$wR(F^2)$ = 0.149
S = 1.03
4399 reflections
205 parameters
All H-atom parameters refined
$\Delta\rho_{max}$ = 0.37 e Å^{-3}
$\Delta\rho_{min}$ = −0.22 e Å^{-3}

Table 1
Selected geometric parameters (Å) for CytFur[a], 5-Furac[b] and Cytosm[c].

	CytFur[a]	5Furac[b]		CytFur[a]	Cytosm[c]
F1—C5	1.3436 (9)	1.348 (1)	O3—C8	1.2517 (10)	1.251 (2)
O1—C2	1.2306 (10)	1.223 (1)	N4—C7	1.3394 (10)	1.341 (2)
O2—C4	1.2331 (10)	1.235 (1)	N4—C8	1.3521 (10)	1.350 (2)
N1—C2	1.3621 (11)	1.362 (1)	N5—C8	1.3753 (11)	1.371 (2)
N1—C6	1.3670 (11)	1.367 (1)	N5—C9	1.3590 (11)	1.353 (2)
N3—C2	1.3763 (11)	1.378 (1)	N6—C7	1.3338 (11)	1.326 (2)
N3—C4	1.3734 (11)	1.374 (1)	C7—C10	1.4255 (12)	1.425 (2)
C4—C5	1.4398 (12)	1.435 (1)	C9—C10	1.3493 (13)	1.333 (2)
C5—C6	1.3347 (12)	1.332 (1)			

Notes: (*a*) this work; (*b*) Hulme *et al.* (2005); (*c*) McClure & Craven (1973).

Table 2
Hydrogen-bond geometry (Å, °).

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
N1—H1···N4^{i}	0.961 (16)	1.837 (16)	2.7970 (11)	176.6 (13)
N3—H3···O2^{ii}	0.925 (16)	1.897 (16)	2.8137 (10)	170.5 (13)
N5—H5···O3^{iii}	0.882 (16)	1.932 (16)	2.8076 (10)	171.9 (13)
N6—H16···O1^{iv}	0.921 (16)	2.141 (16)	2.9103 (11)	140.4 (12)
N6—H161···O1^{v}	0.896 (17)	2.108 (17)	3.0027 (12)	177.6 (14)
C10—H10···O2^{vi}	1.006 (17)	2.421 (17)	3.4202 (11)	172.4 (13)

Symmetry codes: (i) $x-1, y, z$; (ii) $-x, -y+2, -z$; (iii) $-x+1, -y+1, -z+1$; (iv) $-x-1, -y+1, -z$; (v) $x+1, y, z$; (vi) $x-1, y-1, z$.

All H atoms of the molecular adduct were found in a difference Fourier map and were refined isotropically [C—H = 0.976 (15)–1.006 (17) Å and N—H = 0.882 (16)–0.961 (16) Å]. The uncoordinated water molecule is disordered over two sites, O4 and O41, and the occupancies refined to close to 0.5. Eventually, atoms O4 and O41 were refined by imposing that their occupancy factors must add to unity. The attached H atoms could not be reliably located in difference maps and were not included in the final refinement, although they are included in the chemical formula.

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*.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: GD3108). Services for accessing these data are described at the back of the journal.

References

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.

Bernstein, J., Davis, R. E., Shimoni, L. & Chang, N.-L. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 1555–1573.

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.

Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology*, pp. 398–404. Oxford University Press.

Etter, M. C., MacDonald, J. C. & Bernstein, J. (1990). *Acta Cryst.* B**46**, 256–262.

Farrugia, L. J. (1997). *J. Appl. Cryst.* **30**, 565.

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

Hulme, A. T., Price, S. L. & Tocher, D. A. (2005). *J. Am. Chem. Soc.* **127**, 1116–1117.

Hunter, W. N. & Brown, T. (1999). *Oxford Handbook of Nucleic Acid Structure*, edited by S. Neidle, pp. 313–330. Oxford University Press.

McClure, R. J. & Craven, B. M. (1973). *Acta Cryst.* B**29**, 1234–1238.

Motherwell, W. D. S., Shields, G. P. & Allen, F. H. (1999). *Acta Cryst.* B**55**, 1044–1056.

Pinedo, H. M. & Peters, G. F. J. (1988). *J. Clin. Oncol.* **6**, 1653–1664.

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.

Sheldrick, G. M. (1997). *SHELXL97*. University of Göttingen, Germany.

Strazewski, P. & Tamm, C. (1990). *Angew. Chem. Int. Ed. Engl.* **19**, 36–57.

Voet, D. & Rich, A. (1969). *J. Am. Chem. Soc.* **91**, 3069–3075.

Wahl, M. C., Rao, S. T. & Sundaralingam, M. (1996). *Nat. Struct. Biol.* **3**, 24–31.