Acta Chemica Scandinavica, 1999: 53: 57–68
Printed in UK – all rights reserved

Copyright © Acta Chemica Scandinavica 1999

ACTA
CHEMICA SCANDINAVICA
ISSN 0904–213X

The Effect of Hydrogen Bonding on the Structures of Uracil and Some Methyl Derivatives Studied by Experiment and Theory

Gustavo Portalone,[a,*] Luigi Bencivenni,[a,b] Marcello Colapietro,[a] Andrea Pieretti[a,b] and Fabio Ramondo[c,*]

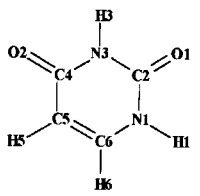
[a]Dipartimento di Chimica, Università 'La Sapienza', I-00185 Roma, Italy, [b]Istituto Nazionale per la Fisica della Materia, via dell'Acciaio 139, I-16153 Genova, Italy and [c]Dipartimento di Chimica, Ingegneria Chimica e Materiali, Università dell'Aquila, I-67100 L'Aquila, Italy

Portalone, G., Bencivenni, L., Colapietro, M., Pieretti, A. and Ramondo, F., 1999. The Effect of Hydrogen Bonding on the Structures of Uracil and Some Methyl Derivatives Studied by Experiment and Theory. Acta Chem. Scand. 53: 57–68. © Acta Chemica Scandinavica 1999.

The equilibrium structures of uracil, thymine (5-methyluracil), 1-methyluracil and 1,3-dimethyluracil have been determined from *ab initio* MO calculations at the HF, MP2 and DFT/6-31G(d) levels. Thymine has also been studied in the solid state by X-ray diffraction. A systematic investigation of the effects of intermolecular association on the structure of uracil and its methyl derivatives has then been carried out by theoretical calculations. Based on information from the solid state, a dimer for 1-methyluracil, a trimer for thymine and a hexamer for uracil are used as theoretical models (at HF and DFT levels) to simulate the neighboring intermolecular interactions found in the crystal. The geometrical perturbations observed in the uracilic moiety, on going from an isolated molecule to the crystal phase, are mainly due to self-association and confirm the fundamental role of conjugative stabilization of the intermolecular hydrogen bonding in such DNA bases.

Owing to the fundamental role of nucleic acids in the functioning and replication of living systems, uracil and its derivatives have been investigated extensively by a variety of experimental and theoretical methods. In particular, methylated derivatives of uracil have been the subject of a vast amount of work, indicating that methylation of DNA is probably involved in various genetic alterations and in the initiation of carcinogenic processes. Furthermore much effort has been devoted to studying the hydrogen bonding intermolecular interactions in these systems since *'hydrogen bonding is the key feature in the biological information transfer mechanisms by the nucleic acids.'*[1]

The structure of the free uracil molecule (Scheme 1)



Scheme 1.

* To whom correspondence should be addressed.

has been studied by gas-phase electron diffraction[2] and by several *ab initio* molecular orbital calculations.[3–8] Regarding methyl derivatives in the gas phase, experimental structural determinations are available from microwave (MW) spectroscopy only for thymine (5-methyluracil),[9] whereas some theoretical studies at different levels of theory have been reported concerning thymine,[4,10–13] 1-methyl uracil,[14] 6-methyluracil[15] and 1,3-dimethyl uracil.[16] Most of the theoretical studies have been devoted to the interpretation of vibrational spectra of molecules isolated in inert gas low-temperature matrices[4,5] or the determination of uracil tautomeric stabilities.[7,10]

The molecular geometries have been determined at the Hartree–Fock (HF) level while the use of electron correlation methods has been limited to uracil and thymine molecules.[7,11–13] Recently, the density functional theory (DFT) approach has been successfully applied in the prediction of the structure and tautomerism of uracil and thymine.[6,7,13]

The crystal structures of uracil[17] and some methyl derivatives[18–23] have been determined by X-ray and neutron diffraction techniques at different levels of accuracy. In particular, the crystal structure of thymine

anhydrate[19] dates back to 1969 and does not provide very accurate results ($R=0.149$, only non-hydrogen atoms located), owing to the poor quality of the photographic data collected. Therefore we thought it would be interesting to carry out a new and more accurate crystal structure determination of the thymine anhydrate.

When bond distances and angles of thymine are compared with those of uracil,[17] 1-methyluracil[18] and 1,3-dimethyluracil[22] it clearly emerges that the uracil skeleton has a structure which is significantly different in the molecules considered. For instance uracil, thymine and 1-methyluracil show different C=O bond lengths whereas in 1,3-dimethyluracil, which is unable to form hydrogen bonds, these bond lengths are very similar. Appreciable discrepancies are also found in ring bond distances moving from uracil to its methyl derivatives. Such structural differences might give us indications of the effect that different intermolecular interactions play on the whole molecular geometry.

However it is extremely difficult to find the origin of such structural differences just on the grounds of such a comparison. In addition to the N–H···O=C hydrogen bonding, the weak C–H···O=C interactions might cause further distortions in the uracil ring.[24–26] Finally, we should bear in mind that the presence of methyl groups is expected to affect the geometry of the uracil ring.[27] Our goal here is to investigate the role of each aforementioned factor in determining the geometry of the whole molecule.

Structural effects caused by methylation of the uracil ring can be analyzed for the isolated molecule by theoretical methods as can the effects of intermolecular association.[28–31] First of all a systematic investigation of the effects of the N–H···O=C and C–H···O=C interactions on the geometry of uracil molecule is reported. Nine cyclic dimers were considered. The analysis of geometrical changes due to dimerization serves as a preliminary approach to a study of the hydrogen bonding in the uracil crystal and its methyl derivatives.

A dimer for 1-methyluracil, a trimer for thymine and a hexamer for uracil are next proposed as suitable models to simulate the neighboring intermolecular interactions in the crystal. The size of such oligomers does not allow the inclusion of electron correlation corrections using conventional *ab initio* methods; however, the DFT methodology can be an alternative and efficient approach to model such large systems.

Experimental

Crystals of thymine (Aldrich) were grown by slow evaporation from an ethanolic solution. A suitable crystal was mounted on Huber CS diffractometer[32] with graphite monochromatized Mo-Kα radiation ($\lambda=0.71069$ Å). Details of the crystal data, data collection, structure solution and refinement are given in Table 1.

All the hydrogen atoms were clearly located through a Fourier synthesis and introduced into the final least-squares refinement with the isotropic temperature factors arbitrarily fixed at 1.2 times the corresponding value of their bonded atoms. Table 2 lists the final fractional coordinates for non-hydrogen atoms. The full CIF (Crystallographic Information File) is available from the authors.

Table 1. Summary of crystal data, data collection, structure solution and refinement for thymine.

(a) Crystal data	
Empirical formula	$C_5H_6N_2O_2$
Molar mass	126.115
Melting point (K)	603
Color, habit	Colorless, prismatic
Crystal size (mm)	$0.3\times0.2\times0.2$
Crystal system	Monoclinic
a/Å	12.889(7)
b/Å	6.852(3)
c/Å	6.784(3)
β/°	104.92(5)
V/Å^3	578.9(5)
Space group	$P2_1/c$
Z	4
$F(0,0,0)$	264.0
D_c	1.447 g cm^{-3}
μ/mm^{-1}	0.114
(b) Data collection[a]	
T/K	298
Unit-cell reflections (θ-range/°)	15 (15–20)
Maximum value of θ/° for reflections	35
hkl range of reflections	−1 20; −1 11; −10 10
Variation in 3 standard reflections	<3%
Reflections measured	3063
Unique reflections	618
Reflections with $\lvert F_o\rvert>2\sigma\lvert F_o\rvert$	604
R of merged reflections	0.027
(c) Structure solution and refinement[b]	
Refinement on	$\lvert F_o\rvert$
Solution method	Direct methods
No. of variables refined	100
Weighting scheme, $1/(a+b\lvert F_o\rvert+c\lvert F_o\rvert^2)$	0.38133, 0.09782, 0.00024
R, R_w, S[c]	0.0458, 0.0637, 1.01
Density range in final Δ-map/e Å^{-3}	−0.27, 0.15(5)
Final shift/error ratio	0.000

[a]Data were collected using ω scan. The intensities were corrected for Lorentz-polarization effects, but not for absorption. [b]All calculations were done on a Pentium PC with SIR CAOS,[47,48] PARST[49] and CRYSTALS[50] packages. Atomic form factors were taken from Ref. 51. [c]$S=\{\Sigma[w(F_o^2-F_c^2)^2]/(n-p)\}^{1/2}$.

Computational methods

All *ab initio* molecular orbital calculations were run on an Alpha AXP-3000/500 cluster of CASPUR c/o CICS at the University of Rome using the GAUSSIAN 94 package.[33]

Table 2. Final fractional coordinates and U_{eq} for the non-hydrogen atoms for thymine.

	x	y	z	U_{eq}/Å^2
O(1)	0.4878(2)	−0.0076(3)	0.2430(4)	0.0564(9)
O(2)	0.7810(2)	−0.3511(3)	0.5833(4)	0.0624(10)
N(1)	0.6387(3)	0.1603(3)	0.4037(4)	0.0493(11)
N(3)	0.6361(3)	−0.1750(3)	0.4166(4)	0.0461(10)
C(2)	0.5824(3)	−0.0071(4)	0.3481(5)	0.0450(11)
C(4)	0.7425(3)	−0.1895(4)	0.5342(6)	0.0476(13)
C(5)	0.7971(3)	−0.0042(4)	0.5895(5)	0.0491(12)
C(6)	0.7436(3)	0.1610(4)	0.5225(6)	0.0493(13)
C(7)	0.9110(4)	−0.0073(6)	0.7179(8)	0.0714(18)

1. Isolated molecules. The geometries of the uracil, 5-methyl-, 1-methyl- and 1,3-dimethyl-uracil isolated molecules were optimized by an analytical gradient-based technique at the Hartree–Fock (HF) level and at the second order of Møller–Plesset perturbation theory (MP2)[34] employing the frozen core approximation (fc). The split valence basis sets 4–31G(d) and 6–31G(d) were used.[35] The structures were constrained to have planar C_s symmetry. For thymine, two models were considered, one with a C–H bond eclipsing the C5=C6 bond (*a* conformation), the other with a C–H eclipsing the C4–C5 bond (*b* conformation). HF/4–31G(d) geometries of monomers and oligomers were compared with the results at the DFT level with regard to the accuracy of the structural changes predicted upon intermolecular association. Geometry optimizations were carried out employing the exchange-correlated functional of Becke[36,37] and Lee *et al.*[38] implemented in the GAUSSIAN 94 program (BLYP).

2. Dimers of uracil. All nine uracil dimers in Fig. 1 consist of two molecules hydrogen bonded pairwise. The six structures [Fig. 1(A)] describe cyclic dimers interacting with two C=O···H–N bonds whereas the remaining ones [Fig. 1(B)] describe two uracil molecules both arranged to form C=O···H–N and C=O···H–C intermolecular bonds. All geometries were fully optimized under C_s or C_{2h} symmetry constraints at the HF/4-31G(d) level.

The interaction energy ΔE was calculated for each dimer using the supermolecular model, $\Delta E = E_{dimer} - 2E_{monomer}$ in the basis set of the complex.

Atomic orbital population analysis was carried out using the natural bond orbital (NBO) method developed by Weinhold[39] with the NBO subroutine of GAUSSIAN 94. Natural atomic charges and p_π atomic orbital occupancy of monomer and dimer were analyzed by MP2/6-31G(d) calculations employing all the active orbitals, MP2(full), at the HF/4-31G(d) optimized geometries.

3. Crystal models. Geometry optimizations were carried out at the HF/4-31G(d) and BLYP/6-31G(d) levels. A centrosymmetric dimer, showing C_{2h} symmetry (Fig. 2), was studied to simulate 1-methyluracil in the crystal. Each monomer was assumed to have a methyl conforma-

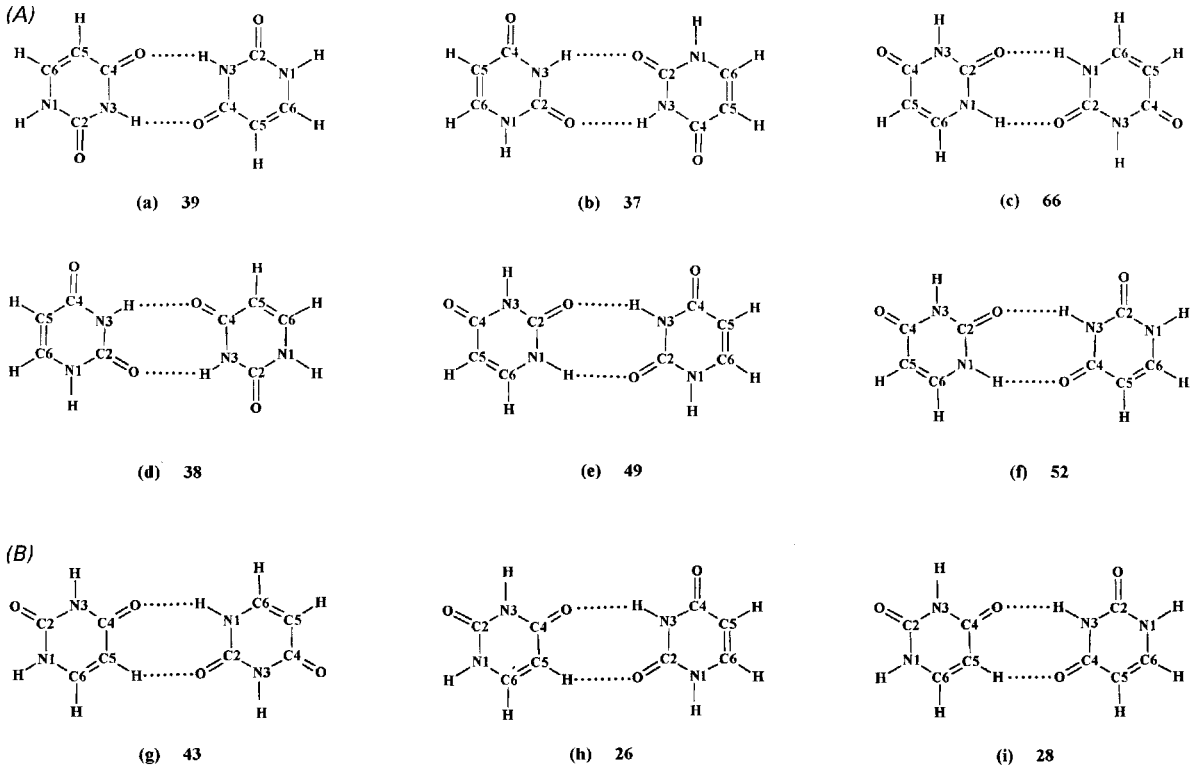


Fig. 1. Cyclic dimers of uracil. The HF/4-31G(d) interaction energy (kJ mol^{-1}) is reported for each structure. The (a–f) pairs (A) show two C=O···H–N hydrogen bonds whereas the (g–i) pairs (B) have both C=O···H–N and C=O···H–C interactions.

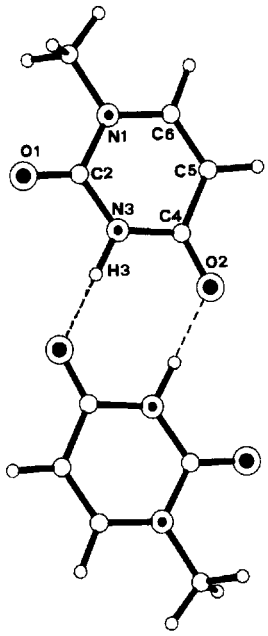


Fig. 2. Centrosymmetric dimer of 1-methyluracil. Oxygen and nitrogen atoms are marked by filled circles. The hydrogen bonds are indicated by the dashed lines.

tion with a C–H bond eclipsing the N1–C6 bond, in agreement with the experimental evidence.[18]

A trimer model reproduced in Fig. 3 was introduced with the aim of simulating the C=O···H–N intermolecular interactions occurring in the thymine crystal structure. The three molecules were assumed to have the methyl conformation found in the crystal (which corresponds to the most stable one at all levels of calculations) whereas the whole trimer was constrained to have planar C_s symmetry.

Both C=O···H–N and C=O···H–C intermolecular interactions occurring in the uracil crystal have been investigated by the hexamer reproduced in Fig. 4. With the aim of reducing the computational efforts, the hexamer was assumed to be planar and centrosymmetric.

Results and discussion

1. Isolated molecules from theoretical calculations. Table 3 gives the geometries of uracil, thymine, 1-methyluracil and 1,3-dimethyluracil determined by calculations at the MP2/6-31G(d) level. A quick scan of the theoretical results clearly indicates that the structure of the uracil moiety is very similar in the molecules here considered. By a detailed analysis we observe that the ring as well as the C=O bond distances of the uracil molecule change to within a few thousands of an ångstrom by methyl substitution. The effect of different orientations of the methyl group on the ring structure of thymine is negligible. Instead, appreciable differences are found in the values of the bond angles, in particular in the region close to the methyl group. Taking as a reference the geometry of the uracil molecule, we find that the bond angle subtending the substituent is 1.7° narrower for thymine, 2.3° for 1-methyluracil and 2.2° (C6–N1–C2) and 2.7° (C2–N3–C4) for 1,3-dimethyluracil. Such changes are consistent with the geometrical deformations caused by the methyl substitution on the benzene ring.[40] Apart from these differences, the theoretical results clearly indicate that the structural features of the uracil

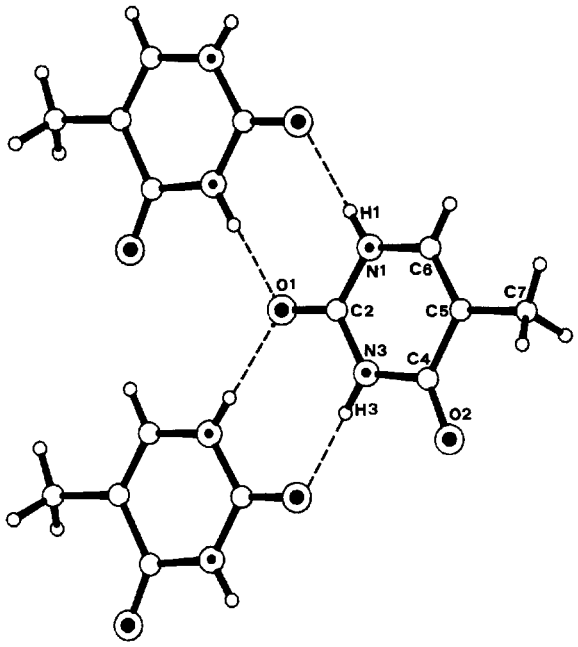


Fig. 3. Theoretical model of thymine in the crystal based on the molecular packing as found in the solid state. Oxygen and nitrogen atoms are marked by filled circles. The hydrogen bonds are indicated by the dashed lines.

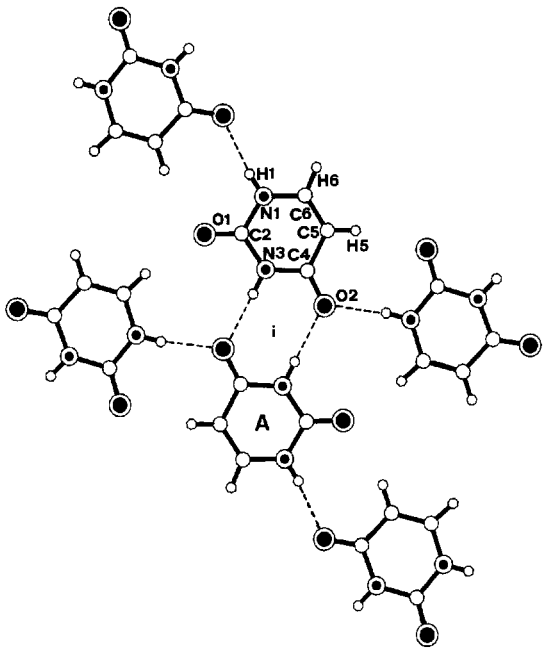


Fig. 4. Theoretical model of uracil in the crystal based on the molecular packing as found in the solid state.[17]

Table 3. Selected MP2/6-31G(d) geometrical parameters of uracil, 5-methyluracil (thymine), 1-methyluracil and 1,3-dimethyluracil molecules (for atomic numbering see Scheme 1).

		5-Methyluracil[a]			
	Uracil	(a)	(b)	1-Methyuracil	1,3-Dimethyluracil
Bond distance (Å)					
N1–C2	1.390	1.386	1.385	1.393	1.393
N3–C2	1.385	1.386	1.386	1.384	1.389
N3–C4	1.408	1.403	1.403	1.406	1.415
C4–C5	1.457	1.462	1.465	1.453	1.451
C5=C6	1.352	1.354	1.356	1.354	1.352
N1–C6	1.377	1.380	1.379	1.377	1.373
C2=O1	1.224	1.225	1.225	1.227	1.231
C4=O2	1.227	1.230	1.230	1.228	1.231
N1–H1	1.013	1.013	1.011		
N3–H3	1.017	1.017	1.017	1.017	
C5–H5	1.082			1.082	1.082
C6–H6	1.085	1.086	1.088	1.086	1.086
C5–C7		1.497	1.503		
N1–Me				1.461	1.461
N3–Me					1.465
Bond angles (°)					
C2–N1–C6	123.9	124.1	123.9	121.6	121.7
N1–C2–N3	112.4	112.2	111.9	113.7	115.6
C2–N3–C4	128.8	128.6	129.1	129.0	126.1
N3–C4–C5	113.2	114.4	114.2	112.6	113.8
C4–C5–C6	119.9	118.4	118.0	120.0	120.6
N1–C6–C5	121.8	122.4	122.9	123.1	122.2
N1–C2=O2	123.2	123.6	123.7	122.5	121.5
N3–C4=O4	120.5	120.7	120.0	120.7	121.4
Me–N1–C2				116.2	116.1
Me–N3–C2					115.3
C7–C5–C4		117.6	118.8		

[a]Rotamers of thymine: (a) has a C–H bond of the methyl group eclipsing the C5=C6 bond, (b) a C–H bond eclipsing the C4–C5 bond.

moiety are not significantly different in the four molecules.

2. Crystal and molecular structure of thymine from X-ray diffraction measurement. A view of the molecular structure of thymine in the crystal is shown in Fig. 5, together with the labelling of the atoms; Table 4 contains the molecular geometry.

Of the several possible tautomeric forms, the diketo tautomer was found in the solid state. These results agree with those found by analysis of the MW spectra in the gas phase[9] and those predicted by *ab initio* MO calculations.[10]

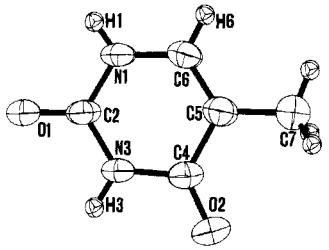


Fig. 5. Drawing of thymine showing the anisotropy of the thermal motion. The thermal ellipsoids are scaled to the 50% probability level.

All the non-hydrogen atoms of the molecule are almost coplanar; the average deviation of the six ring atoms from the least-squares plane is 0.004(4) Å with a maximum deviation of 0.008(4) Å for the C4 atom. Of the three non-hydrogen atoms attached to the ring the O2 atom has the maximum deviation from this plane [0.014(3) Å]. The conformation assumed by the methyl substituent is such that a hydrogen atom is eclipsed with the ring C–C double bond.

The crystal structure of thymine is based on endless ribbons of planar molecules, connected via N–H···O hydrogen bonds (Fig. 6). Within a ribbon, each molecule is hydrogen-bonded to two adjacent molecules; these hydrogen bonds involve the two N–H groups of the ring and the single oxygen atom O1. The N···O separations [2.827–2.833(3) Å] are equal to within experimental error. No shorter interatomic contacts have been found for the second carbonyl oxygen O2, other than 3.387(4) Å, which occurs with the carbon atom C6 of the adjoining molecule related by the unit translation along *b*.

Table 4. Selected experimental geometrical parameters (with esds in parentheses) of uracil, 5-methyluracil (thymine), 1-methyluracil and 1,3-dimethyluracil molecules (for atomic numbering see Scheme 1).

	Uracil[a]	5-Methyluracil[b]	1-Methyluracil[c]	1,3-Dimethyluracil[d]
Bond distance (Å)				
N1–C2	1.371(3)	1.358(4)	1.379(1)	1.374(3)
N3–C2	1.377(2)	1.361(4)	1.377(1)	1.380(3)
N3–C4	1.371(2)	1.401(5)	1.384(1)	1.398(3)
C4–C5	1.430(2)	1.453(4)	1.440(1)	1.436(3)
C5=C6	1.340(2)	1.343(4)	1.357(1)	1.332(3)
N1–C6	1.359(2)	1.384(5)	1.369(1)	1.378(3)
C2=O1	1.215(2)	1.244(4)	1.228(1)	1.225(3)
C4=O2	1.245(2)	1.225(4)	1.241(1)	1.227(3)
N1–Me			1.459(1)	1.459(3)
N3–Me				1.464(3)
C5–C7		1.502(6)		
Bond angles (°)				
C2–N1–C6	122.7(1)	122.5(3)	121.31(5)	120.8
N1–C2–N3	114.0(1)	115.5(2)	115.57(6)	116.7
C2–N3–C4	126.7(2)	126.3(3)	126.54(6)	124.7
N3–C4–C5	115.5(1)	115.1(3)	114.74(6)	114.9
C4–C5–C6	118.9(2)	118.4(3)	119.35(7)	120.4
N1–C6–C5	123.2(2)	122.3(3)	122.48(7)	122.4
N1–C2=O1	123.7(2)	122.4(3)	122.17(6)	121.5
N3–C4=O2	119.2(2)	119.2(3)	119.82(7)	119.6

[a]Ref. 17. [b]This work. [c]Ref. 18. [d]Ref. 22.

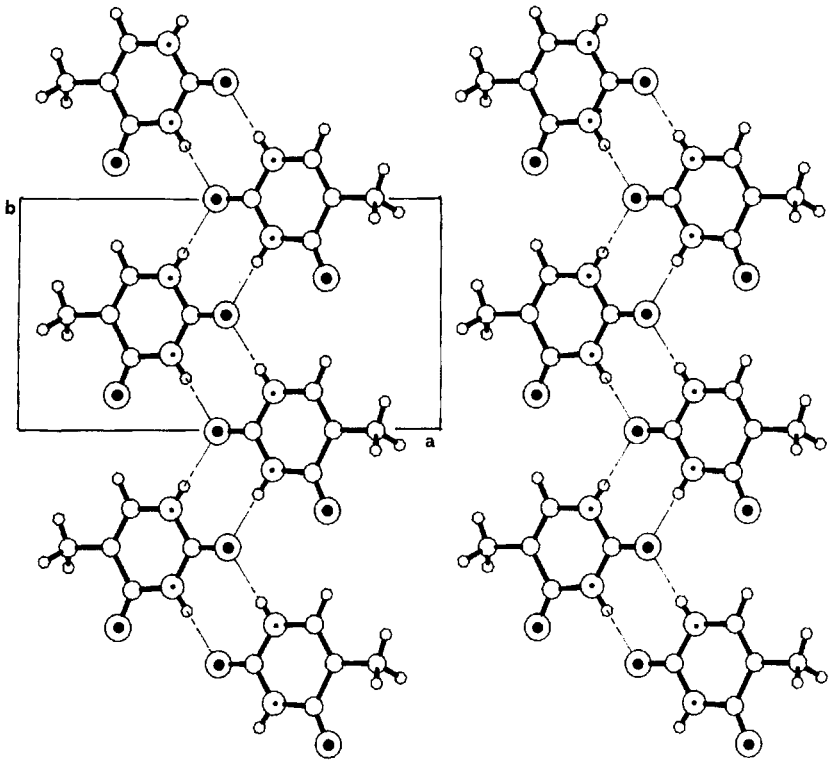


Fig. 6. Projection of the crystal structure down *c* in thymine anhydrate. Oxygen and nitrogen atoms are marked by filled circles. The hydrogen bonds are indicated by the dashed lines.

3. Geometries of isolated and crystal molecules: a comparison. Substantial discrepancies are found when we compare the crystal molecules (see Table 4). Firstly, a significant difference is noted for the carbonyl groups which are almost equivalent in the isolated molecules, but are substantially different in the crystal. For instance, the C4=O2 bond is longer than the C2=O1 bond for uracil [0.030(2) Å] and 1-methyluracil [0.013(1) Å], whereas it is shorter for thymine [0.019(4) Å]. The existence of C=O groups of different lengths is easily explained by observing the respective crystal structures.

In the crystal structure of 1-methyluracil, obtained by a very accurate study[18] by neutron diffraction at 15 K, the molecules form sheets lying in the crystallographic mirror planes of the space group *Ibam*. Within each sheet the molecules form centrosymmetric dimers by N3–H3···O2=C4 [2.806(1) Å] intermolecular hydrogen bonds. Both the carbonyl groups are involved in close interactions with C–H groups.

Similar dimers occur in the crystal structure of uracil,[17] but the N3–H3···O2=C4 [2.865(2) Å] centrosymmetric dimers are connected by additional bifurcated N1–H1···O2=C4 [2.864(2) Å] hydrogen bonds as reproduced in Fig. 4. The remaining carbonyl group is involved in two weaker C–H···O=C interactions (C2=O1···H5–C5 and C2=O1···H6–C6).

The equivalence of the carbonyl groups of 1,3-dimethyluracil is explained by the fact that no C=O···H–N hydrogen bond is possible for such a molecule, although dimers exist in its crystal structure[22] through C=O···H–C interactions. However, the structure of the dimer, studied by AM1 semiempirical methods,[16] is predicted to be very close to that of the monomer. The effect of intermolecular interaction in 1,3-dimethyluracil is very weak indeed, as becomes evident when the geometry of the free molecule is compared with that in the crystal. From the values of the differences between bond distances, $\Delta_1 = r(\text{N3–C4}) - r(\text{N1–C2})$, $\Delta_2 = r(\text{N3–C4}) - r(\text{N3–C2})$, it can be seen that some ring bond distances show a similar pattern in isolated ($\Delta_1 = 0.022$ Å, $\Delta_2 = 0.026$ Å) as well as in crystal ($\Delta_1 = 0.024$ Å, $\Delta_2 = 0.018$ Å) molecules.

In contrast with C=O···H–C interactions, C=O···H–N hydrogen bonding occurring in the uracil, thymine and 1-methyluracil crystals also produces significant deformations in the ring bond distances.

In the thymine isolated molecule, for instance, the N1–C2 and N3–C2 bonds are both shorter ($\Delta_1 = \Delta_2 = 0.017$ Å) than the N3–C4 bond. This is also found in the crystal, although the differences are larger: namely $\Delta_1 = 0.043$ Å and $\Delta_2 = 0.040$ Å. As for thymine, the bond distances of uracil and 1-methyluracil are found to show a different pattern in the gas phase than in the crystal phase. The N1–C2, N3–C2 and N3–C4 bond distances have very similar values for uracil ($\Delta_1 = 0.000$ Å, $\Delta_2 = -0.006$ Å) and 1-methyluracil ($\Delta_1 = 0.005$ Å, $\Delta_2 = 0.007$ Å) in the crystal whereas appreciable differences are noted for uracil ($\Delta_1 = 0.018$ Å, $\Delta_2 = 0.023$ Å) and 1-methyluracil ($\Delta_1 = 0.013$ Å, $\Delta_2 = 0.022$ Å) in the gas phase.

4. The effect of hydrogen bonding from theoretical calculations. Since isolated molecules, having a very similar structure may show noticeable geometrical discrepancies when they are involved in intermolecular interactions, it is worth thoroughly investigating the effect of such interactions on the structure of each molecule by a theoretical approach.

(4a) Uracil dimers. The nine cyclic uracil dimers are shown in Fig. 1. Their geometries were determined at the HF/4-31G(d) level and Table 5 shows the values for the changes of ring bond distances in the uracil ring caused

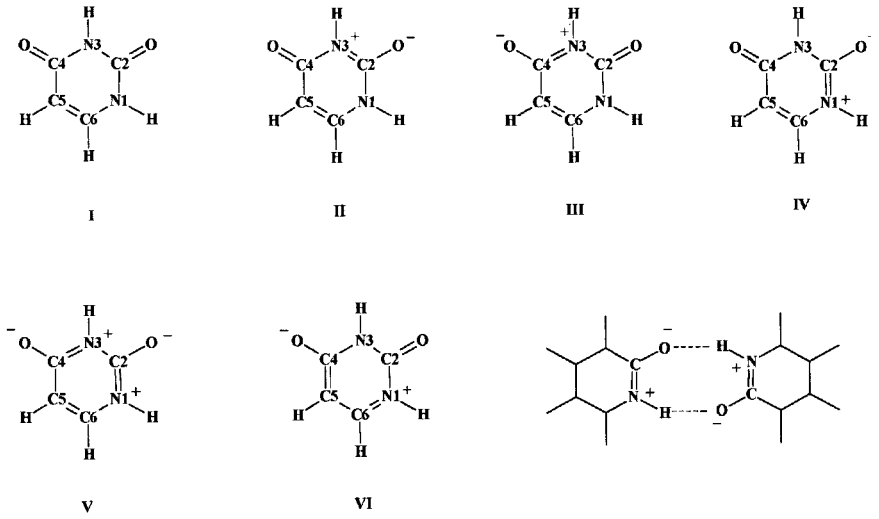
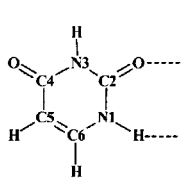
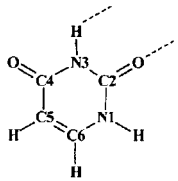
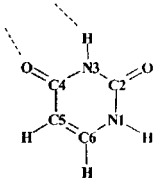
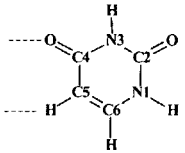
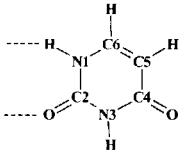
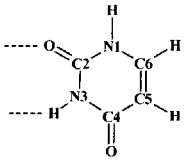
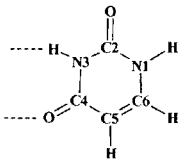
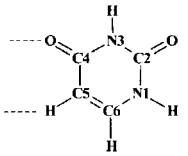


Fig. 7. Canonical forms of uracil (**I–VI**) and the polar canonical form of uracil dimer responsible for stabilization of intermolecular hydrogen bonding.

Table 5. Variations[a] of bond distances (Å) of the uracil molecule calculated at the HF/4-31G(d) level upon cyclic dimerization.

					
N1–C2	−0.015	−0.007	+0.004	−0.003	
N3–C2	−0.007	−0.013	+0.004	+0.005	
N3–C4	+0.004	+0.004	−0.015	−0.011	
C4–C5	−0.004	+0.002	−0.006	−0.007	
C5–C6	+0.002	+0.002	+0.002	+0.002	
N1–C6	0.000	+0.004	−0.005	0.000	
C2=O	+0.015	+0.014	−0.002	−0.001	
C4=O	0.000	−0.002	+0.014	+0.012	
N1–C2	−0.012	−0.005	+0.002	−0.002	
N3–C2	−0.004	−0.010	+0.001	+0.003	
N3–C4	+0.002	0.000	−0.011	−0.007	
C4–C5	−0.003	+0.001	−0.004	−0.005	
C5–C6	+0.002	−0.001	+0.001	+0.001	
N1–C6	−0.002	+0.002	−0.003	+0.001	
C2=O	+0.011	+0.010	−0.001	0.000	
C4=O	+0.001	0.000	+0.009	+0.008	
N1–C2	−0.013	−0.003	+0.004	−0.003	
N3–C2	−0.004	−0.010	+0.001	+0.003	
N3–C4	+0.002	0.000	−0.012	−0.007	
C4–C5	−0.004	+0.001	−0.003	−0.005	
C5–C6	+0.002	0.000	0.000	+0.001	
N1–C6	−0.002	+0.001	−0.003	+0.002	
C2=O	+0.011	+0.010	−0.001	0.000	
C4=O	+0.001	0.000	+0.010	+0.008	
N1–C2	−0.009	−0.001	−0.003		
N3–C2	−0.001	−0.007	+0.008		
N3–C4	+0.001	0.000	−0.009		
C4–C5	−0.004	+0.002	−0.002		
C5–C6	+0.003	−0.002	+0.001		
N1–C6	−0.003	+0.001	−0.003		
C2=O	+0.008	+0.007	−0.001		
C4=O	−0.002	0.000	+0.007		

[a] Calculated as the difference between the bond length of the monomer and the corresponding bond length of the dimer.

by cyclic self-association. The amount of such geometrical variations suggests two important interpretations.

Firstly, C=O···H–N hydrogen bonding causes a concerted distortion of the C=O and C–N bond distances: a lengthening of the C=O bond occurs with simultaneous shortening of the C–N bond distances. Such geometrical variations induced by hydrogen bonding interactions are similar to those found for the amide group of numerous compounds.[41] For example, low-temperature neutron diffraction measurements of a single crystal of urea,[42] *N,N'*-diformohydrazide[43] and acetamide[44] indicate that the C=O bond is significantly longer and the C–N bond

Table 6. Selected bond distances (Å) (with esds in parentheses) of isolated and hydrogen-bonded 1-methyluracil, thymine and uracil.

	HF/4-31G(d)		BLYP-6-31G(d)		Experimental
	Monomer	Dimer	Monomer	Dimer	Neutron diffraction[a]
1-Methyluracil					
N1–C2	1.375	1.379	1.419	1.426	1.379(1)
N3–C2	1.368	1.369	1.397	1.396	1.377(1)
N3–C4	1.388	1.375	1.432	1.412	1.384(1)
C4–C5	1.457	1.453	1.462	1.457	1.440(1)
C5=C6	1.328	1.329	1.365	1.366	1.357(1)
N1–C6	1.372	1.367	1.386	1.381	1.369(1)
C2=O1	1.193	1.192	1.232	1.231	1.228(1)
C4=O2	1.197	1.201	1.233	1.248	1.241(1)
N1–Me	1.456	1.456	1.475	1.475	1.459(1)
N3–H3	0.998	1.008	1.022	1.043	1.043(2)
C5–H5	1.069	1.069	1.089	1.089	1.081(2)
C6–H6	1.072	1.072	1.092	1.093	1.088(2)
Δ_1[b]	0.013	−0.004	0.013	−0.014	0.005
Δ_2[b]	0.020	0.006	0.035	0.016	0.007
	Monomer	Trimer	Monomer	Trimer	X-ray[c]
Thymine					
N1–C2	1.367	1.350	1.407	1.384	1.358(4)
N3–C2	1.369	1.352	1.400	1.377	1.361(4)
N3–C4	1.387	1.390	1.426	1.432	1.401(5)
C4–C5	1.471	1.470	1.476	1.476	1.453(4)
C5–C6	1.328	1.326	1.365	1.365	1.343(4)
N1–C6	1.375	1.377	1.389	1.389	1.384(5)
C2=O1	1.192	1.217	1.230	1.262	1.244(4)
C4=O2	1.193	1.191	1.235	1.233	1.225(4)
N1–H1	0.995	1.004	1.019	1.037	0.82(4)
N3–H3	0.998	1.009	1.022	1.043	0.88(4)
C5–C7	1.508	1.500	1.511	1.510	1.502(6)
C6–H6	1.073	1.072	1.093	1.092	0.98(4)
Δ_1[b]	0.020	0.040	0.019	0.048	0.043
Δ_2[b]	0.018	0.038	0.026	0.055	0.040
	Monomer	Hexamer	Monomer	Hexamer	X-ray[d]
Uracil					
N1–C2	1.372	1.359	1.413	1.397	1.371(3)
N3–C2	1.368	1.368	1.398	1.398	1.377(2)
N3–C4	1.390	1.373	1.433	1.406	1.371(2)
C4–C5	1.461	1.447	1.467	1.449	1.430(2)
C5=C6	1.326	1.332	1.363	1.369	1.340(2)
N1–C6	1.370	1.363	1.385	1.377	1.359(2)
C2=O1	1.191	1.201	1.229	1.241	1.215(2)
C4=O2	1.191	1.213	1.232	1.262	1.245(2)
N1–H1	0.995	1.006	1.019	1.043	0.84(2)
N3–H3	0.995	1.007	1.022	1.041	0.88(2)
C5–H5	1.069	1.070	1.089	1.092	0.93(2)
C6–H6	1.072	1.072	1.092	1.092	0.96(2)
Δ_1[b]	0.018	0.014	0.020	0.009	0.000
Δ_2[b]	0.022	0.005	0.035	0.008	−0.006

[a]Ref. 18. [b]$\Delta_1 = r$(N3–C4) − r(N1–C2); $\Delta_2 = r$(N3–C4) − r(N3–C2). [c]This work. [d]Ref. 17.

Table 7. Selected bond angles (°) (with esds in parentheses) of isolated and hydrogen-bonded 1-methyluracil, thymine and uracil.

	HF/4-31G(d)		BLYP-6-31G(d)		Experimental
	Monomer	Dimer	Monomer	Dimer	Neutron diffraction[a]
1-Methyluracil					
C2–N1–C6	120.9	121.0	121.5	121.5	121.31(5)
N1–C2–N3	114.8	115.1	113.6	114.1	115.57(6)
C2–N3–C4	128.2	127.2	128.9	127.4	126.54(6)
N3–C4–C5	113.1	114.2	112.4	114.3	114.74(6)
C4–C5–C6	119.4	119.1	120.5	119.9	119.35(7)
N1–C6–C5	123.7	123.4	123.3	122.8	122.48(7)
N1–C2=O1	122.4	121.7	122.2	121.4	122.17(6)
N3–C4=O2	120.8	121.4	120.5	121.3	119.82(7)
	Monomer	Trimer	Monomer	Trimer	X-ray[b]
Thymine					
C2–N1–C6	123.5	122.6	124.2	122.5	122.5(3)
N1–C2–N3	113.2	115.6	111.9	115.5	115.5(2)
C2–N3–C4	127.8	126.3	128.5	126.3	126.3(3)
N3–C4–C5	114.8	114.9	114.3	114.5	115.1(3)
C4–C5–C6	117.7	117.8	118.4	118.6	118.4(3)
N1–C6–C5	123.0	122.8	122.6	122.6	122.3(3)
N1–C2=O1	123.3	121.8	123.2	121.3	122.4(3)
N3–C4=O2	120.6	120.8	120.3	120.5	119.2(3)
	Monomer	Hexamer	Monomer	Hexamer	X-ray[c]
Uracil					
C2–N1–C6	123.3	122.7	123.9	122.9	122.7(1)
N1–C2–N3	113.5	114.5	112.3	113.9	114.0(1)
C2–N3–C4	127.9	126.6	128.6	126.4	126.7(2)
N3–C4–C5	113.7	115.4	113.1	115.7	115.5(1)
C4–C5–C6	119.3	118.1	120.3	118.6	118.9(2)
N1–C6–C5	122.2	122.7	121.8	122.4	123.2(2)
N1–C2=O1	122.8	123.0	122.7	122.7	123.7(2)
N3–C4=O2	120.6	119.8	120.3	119.3	119.2(2)

[a]Ref. 18. [b]This work. [c]Ref. 17.

shorter than the corresponding values calculated by *ab initio* methods.[30,45] NBO population analysis of the monomer and dimer indicates that self-association induces significant polarization of the atoms involved in hydrogen bonding. The p_π orbital occupancy reveals that the variations in the charge distribution are mainly due to the shifts in the π-electron density of the C=O and C–N bonds. Therefore the structural changes can be interpreted, in terms of valence bond theory, as an increase in the contribution of polar canonical forms $-C(-O^-)=N-H^+$ relative to the canonical form –C(=O)–N–H. Hence, for cyclic uracil dimers the forms II, III and IV of Fig. 7 play a fundamental role in describing hydrogen bonded molecules and the major contribution comes from the polar form which shows electric charges on those groups forming hydrogen bonding. For instance, the dimer indicated as C in Fig. 1 is successfully described by forms I, IV and, with less importance, by form III. As displayed in Fig. 7, increasing the importance of polar canonical forms such as IV from monomer to dimer makes the C2=O1 ··· H1–N1 hydrogen bonding stronger. The strengthening of intermolecular hydrogen bonding due to resonance is well known[24] and the use of the term 'resonance-assisted hydrogen bond' has been proposed.[46] Significant changes in the C=O and N–C bond distances also occur when uracil forms cyclic dimers by C=O ··· H–N and C=O ··· H–C hydrogen bonds. The structural distortions calculated for (g), (h), (i) dimers [Fig. 1(B)] are again consistent with the increased importance of the canonical forms.

The second important observation can be deduced from an inspection of the results of Table 5. Each column reports the values of the changes in bond distances for an uracil molecule when it interacts with a second molecule through various C=O and H–N couples. Such values indicate that more pronounced geometrical changes occur when the molecule interacts through the H1–N1–C2=O1 framework. In terms of valence bond theory this suggests that structure IV allows the formation of stronger hydrogen bonding. Convincing proof is provided by the fact that more stable complexes (see

Fig. 1) are found when N1–H1 and C2=O1 are involved in intermolecular association and the most stable dimer is the centrosymmetric (c) structure.

(4b) Crystal models. Taking into account that hydrogen bonding causes noticeable geometrical distortions on the uracil skeleton, we next turned our attention to the study of the discrepancies observed between the geometries of isolated and crystal molecules of 1-methyluracil, uracil and thymine.

1-Methyluracil. The geometries of monomer and dimer proposed here to simulate the N3–H3 $\cdots$ O2=C4 hydrogen bonding occurring in the crystal are compared in Tables 6 and 7. Apart from the obvious lengthening of the N–H bond distances, hydrogen bonding causes the concerted shortening of the N3–C4 and lengthening of the C4=O1 bond. Dimerization has a marginal effect on the other bonds, the length of which changes to within 0.005 Å. As already described for uracil dimers, polar canonical forms increase their contribution in the description of the hydrogen bonded molecules. The importance of form III may therefore justify the changes due to self-association.

From Table 6 it appears that the structure of the dimer is in better agreement with the crystal structure than the structure of the free molecule. When the hydrogen bonded model is considered, the values of the differences $\Delta_1 = r(\text{N3–C4}) - r(\text{N1–C2})$ and $\Delta_2 = r(\text{N3–C4}) - r(\text{N3–C2})$ decrease from monomer to dimer and are nearer to the values derived from the crystal structure. Similarly, hydrogen bonding gives rise to a significant difference in the two C=O bond distances, as found for the crystal structure. As regards the bond angles (Table 7), all values calculated at the HF as well as BLYP levels for the dimer are in better agreement with experiment than those of the monomer.

Uracil. The hexamer proposed here takes into account both C=O $\cdots$ H–N and C=O $\cdots$ H–C intermolecular interactions occurring in the crystal. Tables 6 and 7 show the geometry of the monomer, that of the central molecule (A in Fig. 4) of the hexamer along with the experimental X-ray structure.[17] As for 1-methyluracil, the agreement between theoretical and experimental geometries is indeed improved when hydrogen bonding is included. Firstly, the C=O bonds are found to have diffent lengths and, consistent with the experiment, the C4=O2 bond is longer than the C2=O1 bond. Secondly, the central molecule in the hexamer has a ring geometry which is closer to the crystal structure than that of the monomer. Such a structural feature emerges from analysis of the values of Δ_1 and Δ_2 and the values of the ring bond angles.

Finally, a detailed analysis of the changes of each bond length from monomer to hexamer, reveals how hydrogen bonding modifies the ring geometry. The remarkable shortening of the N1–C2, N3–C4 and C4–C5 bond distances suggest the importance of the polar canonical forms V and VI in the description of the crystal molecules of uracil.

Thymine. Tables 6 and 7 report the geometry of the free molecule and that of the central molecule of the trimer. As found for uracil and its 1-methyl derivative, our theoretical simulation of a 'crystal structure', although based on an oversimplified model, significantly improves the agreement between calculated and experimental geometries. When intermolecular interactions are included, the values of the Δ_1 and Δ_2 increase from 0.020/0.018 Å to 0.040/0.038 Å and are nearer to the values derived from the experimental crystal structure (0.043/0.040 Å). In addition, our crystal model reveals that the differences between the two C=O bond distances that are found for the crystal structure are indeed a consequence of intermolecular associations occurring in the crystal phase. A strong change is also observed for the N1–C2–N3 angle (+2.4°) whose value (115.6°) is very close to the experimental one, 115.5(2)°. Better agreement between calculated and experimental bond angles is further found for all ring bond angles. The structural changes calculated for intermolecular association can once again be interpreted, in terms of valence bond theory, by the increase of the polar canonical forms II and IV. On the other hand, II and IV are better proton donors and acceptors than I; increasing their contribution reinforces the N–H $\cdots$ O interaction.

Conclusions

The following conclusions may be drawn from the experimental and theoretical results previously illustrated. (i) The accurate X-ray structure determination of thymine demonstrates that the molecule is present in the crystal as a diketo tautomer. (ii) In the absence of accurate molecular geometries of gas-phase uracil methyl derivatives, structural features of the uracil ring have been obtained from MP2/6-31G(d) optimized geometries. The *ab initio* results indicate that the structure of the uracil ring is very similar in the molecules considered. (iii) Crystal molecules show remarkably different geometries and our theoretical simulations of crystal structures demonstrate that hydrogen bonding is the main factor responsible for the discrepancies observed. (iv) It is worth noting that DFT with BLYP functional gives results that are qualitatively consistent with the HF level as regards the geometrical distortions induced by hydrogen bonding. (v) Our theoretical models confirm the fundamental role of conjugative stabilization of intermolecular hydrogen bonding in such DNA bases.

Acknowledgements. We gratefully acknowledge financial support from MURST.

References

1. Jeffrey, G. A. and Saenger, W. *Hydrogen Bonding in Biological Structures*, Springer-Verlag, Heidelberg, Berlin, New York 1991, p. 235.
2. Ferenczy, F., Hersonyi, L., Rozsondai, B. and Hargittai, I. *J. Mol. Struct. 140* (1986) 71.
3. Brown, R. D., Goodfrey, P. D., McNaughton, D. and Pierlot, A. P. *J. Am. Chem. Soc. 110* (1988) 2329.
4. Les, A., Adamowicz, L., Nowak, M. J. and Lapinski, L. *Spectrochim. Acta, Part A 48* (1992) 1385.
5. Gould, I. R., Vincent, M. A. and Hillier, I. H. *J. Chem. Soc., Perkin Trans. 2* (1992) 69.
6. Estrin, D. A., Paglieri, L. and Corongiu, G. *J. Phys. Chem. 98* (1994) 5653.
7. Gould, I. R., Burton, N. A., Hall, R. J. and Hillier, I. H. *J. Mol. Struct. (Theochem) 331* (1995) 147.
8. Ha, T. K. and Gunthard, H. H. *J. Mol. Struct. (Theochem) 276* (1992) 209.
9. Brown, R. D., Goodfrey, P. D., McNaughton, D. and Pierlot, A. P. *J. Chem Soc., Chem. Commun.* (1989) 37.
10. Ha, T. K. and Gunthard, H. H. *J. Am. Chem. Soc. 115* (1993) 11939.
11. Stewart, E. L., Foley, C. K., Allinger, N. L. and Bowen, J. P. *J. Am. Chem. Soc. 116* (1994) 7282.
12. Sponer, J. and Hobza, P. *J. Phys. Chem. 98* (1994) 3161.
13. Jolibois, F., Cadet, J., Grand, A., Subra, R., Rega, N. and Barone, V. *J. Am. Chem. Soc. 120* (1997) 1864.
14. Piskorz, P. and Wojcik, M. J. *J. Mol. Struct. (Theochem) 332* (1995) 217.
15. Fan, K. and Boggs, J. E. *J. Mol. Struct. (Theochem) 206* (1990) 265.
16. Alcolea Palafiox, M., Iza, N. and Gil, M. *Int. J. Quantum Chem. 51* (1994) 141.
17. Stewart, R. F. and Jensen, L. H. *Acta Crystallogr. 23* (1967) 1102.
18. McMullan, R. K. and Craven, B. M. *Acta Crystallogr., Sect. B 45* (1989) 270.
19. Ozeki, K., Sakaba, N. and Tanaka, J. *Acta Crystallogr., Sect. B 25* (1969) 1038.
20. Gerdil, R. *Acta Crystallogr. 14* (1961) 333.
21. Micklitz, W., Lippert, B., Schollhorn, M. and Thewalt, U. *J. Heterocycl. Chem. 26* (1989) 1499.
22. Banerjee, A., Dattagupta, J. K., Saenger, W. and Rabczenko, A. *Acta Crystallogr., Sect. B 33* (1977) 90.
23. Clowney, L., Jain, S. H., Srinivasan, A. R., Westbrook, J., Olson, W. K. and Berman, H. M. *J. Am. Chem. Soc. 118* (1996) 509.
24. Jeffrey, G. A. *An Introduction to Hydrogen Bonding*, Oxford University Press, 1997, p. 96.
25. Leonard, G. A., McAuley-Hecht, K., Brown, T. and Hunter, W. *Acta Crystallogr., Sect. D 51* (1995) 136.
26. Taylor, R. and Kennard, O. *J. Am. Chem. Soc. 104* (1982) 5063.
27. Mastryukov, V. S., Fan, K. and Boggs, J. E. *J. Mol. Struct. 346* (1995) 173.
28. Ramondo, F., Bencivenni, L., Portalone, G. and Domenicano, A. *Struct. Chem. 5* (1994) 1.
29. Ramondo, F. and Bencivenni, L. *J. Chem. Soc., Perkin Trans. 2* (1994) 1985.
30. Ramondo, F. and Bencivenni, L. *J. Chem. Soc., Perkin Trans. 2* (1995) 1797.
31. Caminiti, R., Pieretti, A., Bencivenni, L., Ramondo, F. and Sanna, N. *J. Phys. Chem. 100* (1996) 10928.
32. Colapietro, M., Cappuccio, G., Marciante, C., Pifferi, A., Spagna, R. and Helliwell, J. R. *J. Appl. Crystallogr. 25* (1992) 192.
33. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. A., Cheeseman, J. R., Keith, T., Petersson, G. A., Montgomery, J. A., Raghavachari, K., Al-Laham, M. A., Zakrzewski, V. G., Ortiz, J. V., Foresman, J. B., Cioslowski, J., Stefanov, B. B., Nanayakkara, A., Challacombe, M., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, J. L., Replogle, E. S., Gomperts, R., Martin, R. L., Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. P., Head-Gordon, M., Gonzalez, C. and Pople, J. A., GAUSSIAN 94; Gaussian Inc., Pittsburgh, PA 1995.
34. Møller, C. and Plesset, M. S. *Phys. Rev. 46* (1934) 618.
35. Poirer, R., Kari, R. and Csizmadia, I. G. *Handbook of Gaussian Basis Sets*, Elsevier, Amsterdam 1985.
36. Becke, A. D. *J. Chem. Phys. 98* (1993) 5648.
37. Becke, A. D. *Phys. Rev. B 38* (1988) 3098.
38. Lee, C., Yang, W. and Parr, R. *Phys. Rev. A 37* (1988) 785.
39. NBO 3.1: Glendening, E. D., Reed, A. E., Carpenter, J. E. and Weinhold, F. Department of Chemistry, University of Winsconsin, Madison.
40. Domenicano, A. In: Hargittai, I. and Hargittai, M., Eds., *Stereochemical Applications of Gas-Phase Electron Diffraction*, VCH, Weinheim 1988, Part B, Chap. 7.
41. Jeffrey, G. A. In: Domenicano, A. and Hargittai, I., Eds., *Accurate Molecular Structures: Their Determination and Importance*, International Union of Crystallography and Oxford University Press, Oxford 1992, Chap. 11.
42. Swaminathan, S., Craven, B. M. and McMullan, R. K. *Acta Crystallogr., Sect. B 40* (1984) 301.
43. Jeffrey, G. A., Ruble, J. R. and McMullan R. K. *Acta Crystallogr., Sect. B 38* (1982) 1508.
44. Jeffrey, G. A., Ruble, J. R., McMullan, R., DeFrees, D. J., Binkley, J. S. and Pople, J. A. *Acta Crystallogr., Sect. B 36* (1980) 2292.
45. Tsuzuki, S. and Tanabe, K. *J. Chem. Soc., Perkin Trans. 2* (1991) 1255.
46. Gilli, P., Ferretti, V., Bertolasi, V. and Gilli, G. In: Hargittai, M. and Hargittai, I., Eds., *Advances in Molecular Structure Research*, JAI Press, Greenwich, CT 1996, Vol. 2, p. 67.
47. Burla, M. C., Camalli, M., Cascarano, G., Giacovazzo, C., Polidori, G., Spagna, R. and Viterbo, D. *J. Appl. Crystallogr. 22* (1989) 389.
48. Camalli, M., Capitani, D., Cascarano, G., Giacovazzo, C. and Spagna, R. *Italian Patent No. 3540c/86*, SIR CAOS User Guide, Istituto di Strutturistica Chimica CNR, Rome, Italy.
49. Nardelli, M. *Comput. Chem. 7* (1983) 95.
50. Watkin, D. J., Carruthers, J. R. and Betteridge, P. W. CRYSTALS User Guide, Chemical Crystallography Laboratory, Oxford, England 1985.
51. *International Tables for X-Ray Crystallography*, Kynoch Press, Birmingham 1974, Vol. 4, p. 71.

Received June 24, 1998.