surements, EPR, NMR, heat capacity...). No doubt that beyond this first example of a new class of 1-D, structurally ordered bimetallic compounds, other derivatives synthesized with dissymmetric bridging ligands such as dithiooxalate anion, will enlarge the zoo of exotic magnetic materials and will allow a deeper insight into the chemistry and physics in low dimension.

Acknowledgment. We thank J. Galy, J. P. Renard, and J. Seiden for helpful discussions and O. Kahn for suggesting this exciting area of investigation.

Registry No. 1, 89958-80-5; **2**, 89958-81-6; **3**, 89958-82-7; **4**, 89958-83-8; **5**, 90025-85-7; $K_2Ni(S_2C_2O_2)_2$, 25360-81-0; $K_2Cu(S_2C_2O_2)_2$, 89958-84-9; $K_2Pd(S_2C_2O_2)_2$, 60490-61-1; $K_2Pt(S_2C_2O_2)_2$, 60490-62-2.

Supplementary Material Available: A listing of thermal parameters for atoms of **1** (Table IX) and **2** (Table X), plane equation coefficients for **1**, **2**, and **4** (Table XI), and calculated and observed structure factors for **1** (Table XII) and **2** (Table XIII) (15 pages). Ordering information is given on any current masthead page.

An ab Initio Study of Tautomerism of Uracil, Thymine, 5-Fluorouracil, and Cytosine

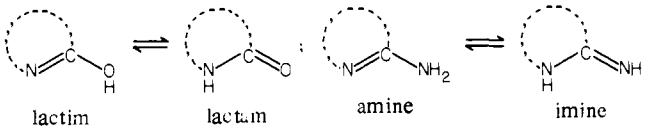
Martin J. Scanlan and Ian H. Hillier*

Contribution from the Department of Chemistry, University of Manchester, Manchester M13 9PL, England. *Received October 25, 1983*

Abstract: Geometry optimization using a 3-21G basis has been employed to study the relative energetics of the six tautomers of uracil, thymine, 5-fluorouracil, and cytosine. These calculations yield molecular geometries in good agreement with available experimental data and correctly predict the most stable tautomer of each species. The relative energetics of three tautomers of uracil are predicted in excellent agreement with experiment. Substitution of uracil at the 5-position by CH_3 or F does not change the order of the stabilities of the tautomers. Our results indicate that the tautomeric equilibria of both uracil and cytosine are sensitive to phase change, and it is suggested that at least two and possibly three tautomers of cytosine may be observed in the gas phase. A simple model of solvation is shown to account for the difference in the order of stability of the tautomers of cytosine and uracil in solution and the gas phase.

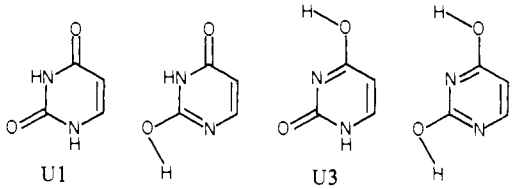
The relative stability of tautomers of the pyrimidine bases uracil, thymine, and cytosine is of fundamental importance to the structure and functioning of nucleic acids. The occurrence of rare tautomers has been put forward as a possible mechanism of spontaneous mutation. Löwdin[1] and more recently Pullman and Pullman[2] have reviewed the subject comprehensively.

We are here concerned with both lactim–lactam and amine–imine tautomerism.

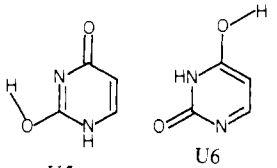


lactim lactam amine imine

Thus six tautomers of uracil and each of its 5-substituted derivatives thymine and 5-fluorouracil

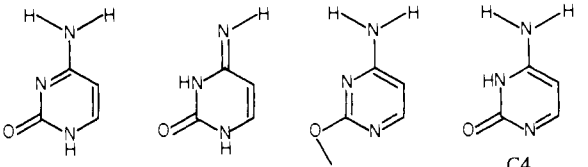


U1 U2 U3 U4

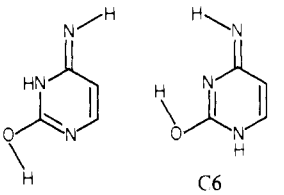


U5 U6

and a further six tautomers of cytosine



C1 C2 C3 C4



C5 C6

will be considered in this paper. The tautomers of uracil and cytosine are denoted U1–U6 and C1–C6 as shown above. The equivalent tautomers of thymine and 5-fluorouracil are U1–U6 with CH_3 or F substituted at C_5 and will be referred to as T1–T6 and FU1–FU6, respectively.

The dominant tautomer for these bases in both the solid and solution has been conclusively determined to be U1, T1, FU1, and C1.[3] Recently, however, Beak[4,5] has drawn attention to the fact that heterocyclic tautomeric equilibria are highly sensitive to environmental effects such as solvent polarity or transition to the

(1) Löwdin, P. O. *Adv. Quantum. Chem.* **1965**, *2*, 213.
(2) Pullman, B.; Pullman, A. *Adv. Heterocycl. Chem.* **1971**, *13*, 77.
(3) Kwiatkowski, J. S.; Pullman, B. *Adv. Heterocycl. Chem.* **1975**, *18*, 199. For a general discussion of tautomerism in heterocycles, see: Katritzky, A. R.; Linda, P. Ed.; "The Tautomerism of Heterocycles"; Academic Press: New York, 1976; Adv. Heterocycl. Chem., Supplement No. 1.
(4) Beak, P.; Fry, F. S.; Lee, J.; Steele, F. *J. Am. Chem. Soc.* **1976**, *98*, 171.
(5) Beak, P. *Acc. Chem. Res.* **1977**, *10*, 186.

0002-7863/84/1506-3737$01.50/0

gas phase. Thus, the equilibrium constant for the 2-pyridone/2-hydroxypyridine equilibrium has been shown to change by a factor of 10^3 on going from a polar to a nonpolar solvent.[5] Hence, it is very likely that the interpretation of data obtained in solution in terms of the relative stability of the tautomers in the gas phase will be erroneous unless the effects of solvation and association are taken into account. However, in the determination of environmental effects a knowledge of the relative gas-phase tautomeric stabilities is an essential prerequisite.

Investigations of the tautomerism of uracil and substituted uracils in the gas phase and in low-temperature argon matrices using UV and IR spectroscopy[6,7] have shown that for the 1-methyl derivatives of uracil, thymine, and 5-fluorouracil the diketo tautomer (U1, T1 and FU1) is the only detectable form. Such definitive evidence has yet to be obtained for cytosine. A study employing photoelectron spectroscopy[8] has shown the lactam–amino tautomer (C1) to be dominant in the gas phase but puts forward the presence of other tautomers as a possible cause of poor spectral resolution.

Thus it is clear that accurate data concerning the stability of the rare tautomeric forms of these bases have yet to be obtained. Furthermore, it has been pointed out by Shugar[7] that since spectroscopic methods are incapable of the detection of less than 0.1–1.0% of a minor tautomer such data are unlikely to be obtained directly by these methods. The frequency of spontaneous mutation, however, is in the range 10^{-8}–10^{-11} per base pair replication[9] and could therefore be readily explained by the presence of a tautomeric form that only exists in concentrations several orders of magnitude lower than those observable by direct experiment. The enthalpy difference between U1, U3, and U4 has recently been estimated experimentally by the direct extrapolation of the energy difference between methyltropic isomers to that of the corresponding protomers.[10] However, due to the large errors associated with this method the results obtained must be viewed with caution.

In an attempt to predict accurate energy differences between the various tautomers numerous quantum mechanical studies have been undertaken.[11-24] The inconsistency of the results produced and, indeed, occasionally their complete disagreement with established experimental results have led to increasing scepticism as to the value of such calculations,[4,5,7,10] and their ability to make a useful contribution to the determination of tautomeric equilibria. Early CNDO/2 calculations for uracil, thymine,[12,13] and their 1-methyl derivatives[14] all predicted the most stable tautomer correctly despite large discrepancies in the tautomeric energy differences. CNDO/2 calculations for cytosine[12,13,15] correctly predict the most stable tautomer but disagree completely as to its relative stability. The first estimate of the tautomeric equilibria of cytosine and thymine by ab initio methods[16] predicted C3 and T4 to be the most stable in disagreement with experiment. These calculations have all been impaired by the lack of molecular geometries for the rare tautomers, requiring them to be estimated from known structural data. Such estimation probably accounts in part at least for the discrepancy in the CNDO/2 results.

More recently, algorithms have become available that include efficient optimization of the molecular geometry thus reducing the problem of determining a structure for the rare tautomers. Nevertheless, the application of MINDO/2[17] including geometry optimization to the tautomers of uracil, thymine, and cytosine predicted the lactim tautomers (U3 and T3) to be the most stable. MINDO/3[18] predicts the most stable tautomer correctly for uracil and cytosine; however, the use of semiempirical geometries in ab initio calculations may lead to a complete reversal of the relative stabilities predicted at the semiempirical level.[19] Ab initio results including geometry optimization in a minimal basis correctly predict U1 and T1 to be more stable than U3 and T3[20,21] while for cytosine C2 is predicted to be marginally more stable than C1.[22] The effect of using geometries obtained by different methods in single-point calculations with a variety of wave functions has been discussed by Mondragon[23] and serves to highlight the inconsistency of results obtained to date. A recent MNDO study of tautomerism in uracil, thymine, and cytosine incorrectly predicts a lactim tautomer to be the most stable in each case.[24]

Of the methods discussed above it is apparent that only a few if any of the results reported provide an accurate estimate of the relative stability of tautomers of the pyrimidine bases. To assess the accuracy of the different approaches it is essential that they are applied to a gas-phase tautomeric equilibrium for which there are accurate experimental data available. For the 2-pyridone/2-hydroxypyridine equilibrium such data are available,[4,25] and a number of these methods have been utilized in attempts to study this equilibrium.[26] All the methods tested in this manner, CNDO/2, MINDO/2, MINDO/3, MNDO, and ab initio (STO-3G), overestimate the stability of the lactim tautomer by between 12 kJ mol^{-1} (MINDO/3) and 140 kJ mol^{-1} (CNDO/2). The most accurate, MINDO/3, is also found to predict the most stable tautomer of uracil and cytosine correctly.

The most recent attempts to predict the 2-pyridone/2-hydroxypyridine equilibrium by ab initio methods[27-29] used complete geometry optimization in a split valence 3-21G basis set and yielded results accurate to within 10 kJ mol^{-1}. The utilization of geometries optimized at a 3-21G level in calculations employing extended basis sets including polarization functions[29] gave results within 2 kJ mol^{-1} of the experimental enthalpy of tautomerization.

The purpose of the present study is to extend the application of these ab initio methods of proven accuracy to the determination of the relative stability of tautomers of the pyrimidine bases. It is expected that such calculations will predict the enthalpy difference between the tautomers to within 10 kJ mol^{-1} and will give the correct order of stability for those tautomers whose energy differences are greater than this value. The molecular geometries determined at this level are the most accurate that it is currently feasible to compute and as such will provide an essential starting point for calculations employing basis sets that include polarization functions and for the study of correlation effects.

We have included 5-fluorouracil in this study because since its introduction as a base analogue by Heidelberger in 1957[30] it has

(6) Nowak, M. J.; Szczepaniak, K.; Barski, A.; Shugar, D. *Z. Naturforsch., C: Biosci.* **1978**, *33C*, 876.
(7) Shugar, D.; Szczepaniak, K. *Int. J. Quantum Chem.* **1981**, *20*, 573.
(8) Yu. C.; Peng, S.; Akiyama, I.; Lin, J.; Le Breton, P. R. *J. Am. Chem. Soc.* **1978**, *100*, 2303.
(9) Freese, E. *J. Theor. Biol.* **1962**, *3*, 82. Drake, J. W. *Nature (London)* **1969**, *221*, 1132.
(10) Beak, P.; White, J. M. *J. Am. Chem. Soc.* **1982**, *104*, 7073.
(11) Bodor, N.; Dewar, M. J. S.; Harget, A. J. *J. Am. Chem. Soc.* **1970**, *92*, 2929.
(12) Fujita, H.; Imamura, A.; Nagata, C. *Bull. Chem. Soc. Jpn.* **1969**, *42*, 1467.
(13) Sorarrain, O. M.; Castro, E. A. *Chem. Phys. Lett.* **1973**, *19*, 422.
(14) Abdulnur, S. *J. Theor. Biol.* **1976**, *58*, 165.
(15) Breen, D. L.; Flurry, R. L., Jr. *Theor. Chim. Acta* **1971**, *23*, 138.
(16) Goddard, J. D.; Mezey, P. G.; Csizmadia, I. G. *Theor. Chim. Acta* **1975**, *39*, 1.
(17) Zielinski, T. J.; Rein, R. *Int. J. Quantum Chem.* **1978**, *14*, 851.
(18) Czerminski, R.; Lesyng, B.; Pohorille, A. *Int. J. Quantum Chem.* **1979**, *16*, 605.
(19) Zielinski, T. J.; Shibata, M.; Rein, R. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1979**, *6*, 475.
(20) Zielinski, T. J.; Shibata, M.; and Rein, R. *Int. J. Quantum Chem.* **1981**, *19*, 171.
(21) Zielinski, T. J. *Int. J. Quantum Chem.* **1982**, *22*, 639.
(22) Palmer, M. H.; Wheeler, J. R.; Kwiatkowski, J. S.; Lesyng, B. *J. Mol. Struct. (Theochem.)* **1983**, *92*, 283.
(23) Mondragon, A.; Ortega Blake, I. *Int. J. Quantum Chem.* **1982**, *22*, 89.
(24) Buda, A.; Sygula, A. *J. Mol. Struct. (Theochem.)* **1983**, *92*, 255.
(25) Cook, M. J.; El-Abbady, S.; Katritzky, A. R.; Guimon, C.; Pfister-Guillouzo, G. *J. Chem. Soc., Perkin Trans. 2* **1977**, 1652. Guimon, C.; Garrabe, G.; Pfister-Guillouzo, G. *Tetrahedron Lett.* **1979**, *28*, 2585. Brown, R. S.; Tse, A.; Vederas, J. C. *J. Am. Chem. Soc.* **1980**, *102*, 1174.
(26) Krebs, C.; Hofmann, H. J.; Kohler, H. J.; Weiss, C. *Chem. Phys. Lett.* **1980**, *69*, 537.
(27) Scanlan, M. J.; Hillier, I. H.; Davies, R. H. *J. Chem. Soc., Chem. Commun.* **1982**, 685.
(28) Schlegel, H. B.; Gund, P.; Fluder, E. M. *J. Am. Chem. Soc.* **1982**, *104*, 5347.
(29) Scanlan, M. J.; Hillier, I. H.; MacDowell, A. A. *J. Am. Chem. Soc.* **1983**, *105*, 3568.

become of great interest in clinical and experimental cancer chemotherapy.[31] The molecular consequences of incorporation of 5-fluorouracil into RNA are numerous and have been discussed in detail by Mandel.[31] Among them is the tendency of 5-fluorouracil to cause mutation. This has been attributed to the increased stability of the rare lactim tautomer (FU3) relative to the equivalent tautomer of uracil.[14,31] CNDO/2[12,14] and more recently ab initio methods[21] have been used to estimate the stability of FU3.

Computational Details

Except for pyrimidine, which has been optimized in both an STO-3G[32] and a 3-21G basis, all geometry optimizations have been performed in the 3-21G split-valence basis set of Binkley et al.,[33] assuming molecular planarity. Convergence was considered to have been achieved when the largest derivative of the energy with respect to any of the nuclear coordinates was less than 0.001 hartree/bohr. Bond lengths and angles are then estimated to have converged to within 0.001 Å and 0.1°, respectively. All calculations were performed with use of the program GAMESS,[34] modified to include the rapid evaluation of the first derivatives of the two-electron integrals as described by Schlegel[35] and implemented on the CDC 7600 of the University of Manchester Regional Computer Centre.

The optimized geometries of the uracil tautomers (U1–U6) were used as starting points for the optimization of thymine and 5-fluorouracil. Complete optimization of the tautomers T1, T2, T3, FU1, FU2, and FU3 indicated that the effect of substitution upon the relative energetics of the tautomers of uracil was small. It was therefore considered quite adequate to perform only a partial geometry optimization of the remaining less stable tautomers (T4, T5, T6, FU4, FU5, and FU6) giving relative energies to within 4 kJ mol^{-1} of the completely optimized structures.

Previous calculations[16,17,24,36] for T1 have shown inconsistency in the orientation of the methyl group with respect to the oxygen atom at the C_4 position, as to whether it should be staggered[24] or eclipsed[17,36] relative to the C_4–C_5 bond. To resolve this discrepancy T1 has been completely optimized with the methyl group in both the staggered and eclipsed conformation. The staggered conformation was found to be more stable by ~4.8 kJ mol^{-1} and has therefore been used in all of the other thymine tautomers.

In tautomers U6, T6, and FU6 the O–H group may be either cis or trans with respect to the N_3–C_4 bond. Schlegel[37] has shown that for 2-hydroxy-4-pyridone the most stable conformation corresponds to the O–H group trans to the N–C bond, the cis–trans energy difference being ~30 kJ mol^{-1}. The trans conformation is therefore employed in U6. Substitution at C_5, as in T6 and FU6, leads to increased steric interaction in the trans conformation. Partial optimization of T6 and FU6 in both the cis and trans conformations has therefore been performed and shows the trans conformation to be the most stable by 42 kJ mol^{-1} for FU6 and by 20 kJ mol^{-1} for T6.

In the case of the cytosine tautomer C3 complete optimization has been performed with the O–H bond both cis and trans with respect to the N_1–C_2 bond to determine the most stable conformation. These conformations are referred to as A and B, respectively. A is found to be more stable by 3.4 kJ mol^{-1}.

For reasons of computational economy ab initio calculations have previously been carried out at molecular geometries derived from semiempirical methods to obtain tautomeric energy differences. To assess the reliability of such an approach we have used MNDO geometries[24] together with a 3-21G basis to calculate the energetics of C1–C6. Significant errors in the MNDO geometries are apparent and thus corrections were made in the C–H and N–H bond lengths of the MNDO structures by modifying them to 1.07 and 1.0 Å, respectively, and making the amine groups planar. The energies obtained by using these corrected MNDO geometries in a 3-21G basis are denoted MNDO//3-21G. Using uncorrected STO-3G geometries[21] for U1 and U3 we have also obtained STO-3G//3-21G results for these tautomers.

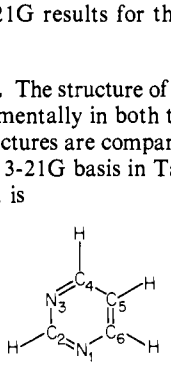
Table I. Comparison of the STO-3G, 3-21G, and Experimental Structures for Pyrimidine[a]

	STO-3G	3-21G	exptl gas[38]	exptl crystal[39]
	Bond Length			
N_1C_2	1.354	1.329	1.340	1.338
N_1C_6	1.354	1.332	1.340	1.340
C_5C_6	1.386	1.382	1.393	1.391
C_2H	1.089	1.067	1.099	1.01
C_5H	1.081	1.069	1.099	0.92
C_6H	1.088	1.070	1.099	0.97
mean deviation[b]	0.012	0.010		
	Bond Angle			
$N_1C_2N_3$	128.0	124.6	127.6	126.8
$C_2N_1C_6$	114.6	117.7	115.5	116.1
$N_1C_6C_5$	122.6	121.5	122.3	122.2
$C_4C_5C_6$	117.5	116.9	116.8	116.6
N_1C_2H	116.0	117.7	116.2	115
N_1C_6H	116.5	117.0	115.3	113
C_6C_5H	121.3	121.6	121.6	125
mean deviation[b]	0.6	1.5		

[a] In all tables bond lengths and angles are given in Å and deg, respectively. E(STO-3G) = −259.38745 au, E(3-21G) = −261.20617 au. [b] With respect to gas-phase geometry,[38] for heavy atoms only.

Results and Discussion

Molecular Geometries. The structure of pyrimidine has recently been determined experimentally in both the gas phase[38] and the solid state.[39] These structures are compared with those predicted from an STO-3G and a 3-21G basis in Table I. The ring numbering employed herein is



with groups substituted at C_2 and C_4 numbered 7 and 8, respectively.

The agreement between the theoretical structures predicted in both basis sets and the experimental geometry is very good for the ring atoms. If we are considering tautomerism the ability of the basis set to predict the position of hydrogen atoms is important, and it is clear from Table I that there is a consistent discrepancy of 0.03 Å between the experimental gas-phase C–H bond lengths and those predicted in a 3-21G basis. This is probably due in large measure to the experimental gas-phase C–H bond lengths being uniformly too long due to vibrational effects. The 3-21G basis has previously been shown to reproduce the C–H bond lengths of a variety of small molecules to within 0.02 Å[33] and those of pyridine to within 0.015 Å.[29] In view of these facts the observed discrepancy is not considered to be significant.

The optimized geometries obtained for the tautomers are given in Figures 1–4. The geometries for the partially optimized

(30) Heidelberger, C.; Chaudhuri, N. K.; Danneberg, P.; Mooren, D.; Griesbach, C.; Duschinsky, R.; Schnitzer, R. J.; Pleven, E.; Scheiner, J. *Nature (London)* **1957**, *179*, 663.
(31) Mandel, H. G. *Prog. Mol. Subcell. Biol.* **1969**, *1*, 82.
(32) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.
(33) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.
(34) Dupuis, M.; Spangler, D.; Wendoloski, J. J. "GAMESS NRCC Software Catalog", Vol. 1, Program No. 2GO1, 1980.
(35) Schlegel, H. B. *J. Chem. Phys.* **1982**, *77*, 3676.
(36) Del Bene, J. E. *J. Phys. Chem.* **1983**, *87*, 367.
(37) Schlegel, H. B. *Int. J. Quantum Chem.* **1982**, *22*, 1041.
(38) Fernholt, L.; Romming, C. *Acta Chem. Scand., Ser. A* **1978**, *A32*, 271.
(39) Furberg, S.; Grogaard, J.; Smedsrud, B. *Acta Chem. Scand., Ser. B* 1979, *B33*, 715.

Table II. Bond Lengths of Uracil (U1)

bond length	3-21G[a]	STO-3G[21]	MINDO/2[17]	MINDO/3[23]	MNDO[24]	exptl[b,40]
N_1C_2	1.380	1.427	1.401	1.391	1.416	1.374 (0.019)
C_2N_3	1.374	1.419	1.399	1.380	1.403	1.381 (0.022)
N_3C_4	1.396	1.432	1.430	1.401	1.429	1.380 (0.013)
C_4C_5	1.457	1.492	1.455	1.469	1.477	1.444 (0.024)
C_5C_6	1.326	1.323	1.350	1.359	1.361	1.343 (0.026)
C_6N_1	1.375	1.405	1.411		1.396	1.370 (0.022)
C_2O_7	1.211	1.219	1.219	1.215	1.227	1.219 (0.020)
C_4O_8	1.212	1.221	1.222	1.211	1.225	1.233 (0.023)
N_1H	0.998	1.019	1.118	1.028	1.002	
N_3H	1.002	1.019	1.121	1.034	1.006	
C_5H	1.066	1.072	1.181	1.102	1.089	
C_6H	1.069	1.086	1.187	1.113	1.095	
mean deviation	0.012	0.032	0.021	0.015	0.026	

[a] This work. [b] Standard deviations are given in parentheses.

Table III. Bond Angles of Uracil (U1)

bond angle	3-21G[a]	STO-3G[21]	MINDO/2[17]	MINDO/3[23]	MNDO[24]	exptl[b,40]
$N_1C_2N_3$	113.3	112.7	111.7	109.5	115.2	115.4 (1.5)
$C_2N_3C_4$	128.1	127.9	130.1	131.7	125.2	126.4 (1.4)
$N_3C_4C_5$	113.5	112.6	111.6	111.4	114.8	114.1 (1.6)
$C_4C_5C_6$	119.8	121.2	122.1	120.8	121.0	120.7 (1.9)
$C_5C_6N_1$	122.4	122.6	120.7		120.5	121.2 (1.2)
$C_6N_1C_2$	122.9	122.7	123.8		123.3	122.0 (1.4)
$N_1C_2O_7$	122.9	123.8	122.1	123.1	121.3	122.9 (1.3)
$N_3C_4O_8$	120.7	120.3	120.8	120.3	117.6	120.5 (1.5)
C_6N_1H	120.8	120.8	117.5		118.4	
C_2N_3H	115.6	115.4	113.9	112.8		
C_4C_5H	118.4	117.4	119.5	119.4	117.6	
C_5C_6H	121.9	121.9	124.7	123.8	118.1	
mean deviation	1.0	1.2	1.8	2.4	1.1	

[a] This work. [b] Standard deviations are given in parentheses.

Table IV. Bond Lengths of Cytosine (C1)

bond length[a]	3-21G[b]	minimal basis[12]	MINDO/2[17]	MINDO/3[22]	MNDO[4]	exptl[c,40]
N_1C_2	1.415	1.458	1.440	1.416	1.438	1.392 (0.015)
C_2N_3	1.369	1.441	1.372	1.354	1.404	1.358 (0.013)
N_3C_4	1.298	1.344	1.320	1.324	1.329	1.339 (0.007)
C_4C_5	1.443	1.489	1.442	1.456	1.458	1.433 (0.015)
C_5C_6	1.337	1.356	1.357	1.376	1.327	1.357 (0.026)
C_6N_1	1.354	1.416	1.409	1.355	1.383	1.360 (0.008)
C_2O_7	1.211	1.293	1.211	1.217	1.224	1.237 (0.024)
C_4N_8	1.344	1.413	1.378	1.335	1.402	1.324 (0.020)
N_1H	0.998	1.024	1.121	1.033	1.002	
N_8H_C	0.998	1.022	1.109	1.014	1.003	
N_8H_T	0.995	1.019	1.116	1.015	1.005	
C_5H	1.067	1.076	1.190	1.102	1.087	
C_6H	1.070	1.089	1.195	1.116	1.094	
mean deviation	0.020	0.052	0.027	0.015	0.034	

[a] H_C and H_T refer to the hydrogens attached to N_8 that are cis and trans respectively with respect to N_3. [b] This work. [c] Standard deviations are given in parentheses.

tautomers are not included (T4–T6,FU4–FU6).

To assess the accuracy of the geometries reported here our optimal bond lengths and angles of U1 and C1 are compared in Tables II–V with those previously determined by other theoretical calculations and with the average ring structures obtained from X-ray studies.[40]

The structures of the pyrimidine bases in the crystal lattice are often influenced by hydrogen bonding so that agreement between the theoretical and the average crystallographic structures is not as good as that found when either the crystal structure is not affected by hydrogen bonding or when comparison is made with gas-phase data. In the absence of hydrogen bonding in the solid state, for instance in the case of pyrimidine, the gas-phase and crystal structure of the ring are extremely close (Table I).

The 3-21G geometries consistently give the smallest mean absolute deviation from experiment except for the cytosine bond lengths where MINDO/3 results show a smaller deviation. However, the bond angles of cytosine predicted by MINDO/3 give the largest deviation from experiment. Generally, MINDO/3 gives good bond lengths and poor bond angles while MNDO and the minimal basis sets give poor bond lengths and good bond angles. The largest mean absolute deviation from experiment is 0.052 Å for the bond lengths of cytosine predicted by a 7s3p/3s basis contracted to minimal.[22] An STO-3G basis gives a deviation of 0.043 Å for the same bond lengths.[41] The most important observation to be made in considering deviation from experimental results is that the 3-21G geometries are the only ones that predict both bond lengths *and* bond angles with reasonable accuracy in all cases considered.

(40) Voet, D.; Rich, A. *Prog. Nucleic Acid Res. Mol. Biol.* **1970**, *10*, 196.

(41) Del Bene, J. E. *J. Comput. Chem.* **1983**, *4*, 226.

Table V. Bond angles of Cytosine (C1)

bond angle[a]	3-21G[b]	minimal basis[22]	MINDO/2[17]	MINDO/3[22]	MNDO[24]	exptl[c 40]
$N_1C_2N_3$	115.2	117.3	112.6	112.8	117.4	118.6 (1.3)
$C_2N_3C_4$	122.1	118.2	124.0	124.8	119.9	120.5 (1.3)
$N_3C_4C_5$	122.8	124.9	123.9		123.2	121.5 (1.7)
$C_4C_5C_6$	116.1	117.1	115.5		117.8	117.0 (2.0)
$C_5C_6N_1$	120.8	120.1	119.6	118.7	119.2	121.2 (0.8)
$C_6N_1C_2$	123.0	122.3	124.4	126.1	122.5	121.2 (1.2)
$N_1C_2O_7$	118.9	118.2	119.7	120.6	118.7	118.9 (1.2)
$N_3C_4N_8$	118.3	117.2	117.5	115.5	116.2	118.3 (1.6)
C_6N_1H	121.3	121.4	116.0	114.5	118.1	
$C_4N_8H_C$	117.9	117.8	120.3	126.0	114.1	
$C_4N_8H_T$	122.5	120.3	119.3	123.6	115.7	
C_4C_5H	121.7	120.0	122.9		121.2	
C_5C_6H	122.4	123.5	122.1	121.4	123.6	
mean deviation	1.2	1.4	2.5	3.7	1.2	

[a] H_C and H_T refer to the hydrogens attached to N_8 that are cis and trans respectively with respect to N_3. [b] This work. [c] Standard deviations are given in parentheses.

Table VI. Total Energies (au) for the Tautomers of Uracil, Thymine, 5-Fluorouracil, and Cytosine

	tautomer					
	1	2	3	4	5	6
uracil	−410.163 10	−410.135 63	−410.131 95	−410.124 59	−410.121 45	−410.117 54[b]
thymine	−448.987 82	−448.960 76	−448.954 79	−448.947 52	−448.946 99	−448.935 82[b]
5-fluorouracil	−508.470 20	−508.444 98	−508.438 28	−508.433 14	−508.431 55	−508.426 46[b]
cytosine	−390.416 17	−390.415 52	−390.410 11 (A)[a] −390.408 80 (B)	−390.404 81	−390.387 13	−390.380 07

[a] In A and B the hydroxyl hydrogen is cis and trans respectively to the N_1–C_2 bond. [b] The hydroxyl hydrogen is trans with respect to the N_3–C_4 bond.

Table VII. Relative Energies (kJ mol^{-1}) of Uracil Tautomers

	tautomer					
method	U1	U2	U3	U4	U5	U6
CNDO/2[a 12]	0.0		43.4			
CNDO/2[a,c 14]	0.0		86.4			
Hückel[a 11]	0.0	74.8	124.6	139.9		99.7
MINDO/2[17]	0.0		−59.4			
MINDO/3[18]	0.0	41.0	21.3	55.6	58.2	33.0
MNDO[24]	0.0	−5.4	3.0	−35.9	33.1	30.8
MINDO/2//STO-3G[19]	0.0		106.7			
STO-3G[21]	0.0		27.7			
ab initio[23]	0.0		104.2			
STO-3G//3-21G[b]	0.0		92.8			
3-21G[b]	0.0	72.1	81.8	101.1	109.3	119.6
exptl[10]	0.0		79.5 (±25.1)	92.0 (±41.8)		

[a] Optimized geometries have not been used. [b] This work. [c] 1-Methyluracil.

Table VIII. Relative Energies (kJ mol^{-1}) of Thymine Tautomers

	tautomer					
method	T1	T2	T3	T4	T5	T6
CNDO/2[a 12]	0.0		49.2			
CNDO/2[a 13]	0.0		4.4			
CNDO/2[a,c 14]	0.0		93.1			
MINDO/2[17]	0.0		−37.7			
STO-3G[a 16]	0.0	14.6	17.9	−62.9	56.0	64.2
STO-3G[21]	0.0		28.2			
MNDO[24]	0.0	−6.6	2.6	−37.3	32.0	30.6
3-21G[b]	0.0	71.0	86.7	105.8	107.2	136.5

[a] Optimized geometries have not been used. [b] This work. [c] 1-Methylthymine.

Table IX. Relative Energies (kJ mol^{-1}) of 5-Fluorouracil Tautomers

	tautomer					
method	FU1	FU2	FU3	FU4	FU5	FU6
CNDO/2[a 12]	0.0		41.5			
CNDO/2[a,c 14]	0.0		82.1			
STO-3G[21]	0.0		31.9			
3-21G[b]	0.0	66.2	83.8	97.3	101.5	114.8

[a] Optimized geometries have not been used. [b] This work. [c] 1-Methyl-5-fluorouracil.

The changes in molecular geometry associated with lactam–lactim tautomerism are similar in all four bases. The most significant of these changes are in the C–O bond length (0.13 Å), the C–N bond length (0.1 Å), the internal angle at the carbonyl carbon (10°), and a corresponding change (5°) in the two adjacent angles. Changes in the other bond lengths and angles are typically an order of magnitude smaller than these values (see Figures 1–4).

Amine–imine tautomerism involves a change in both C–N bond lengths involved by ~0.1 Å, and the internal angle at which the amine group is substituted varies by ~10°.

The effects of substitution upon the geometry of uracil are very small in agreement with experiment.[40] Upon substitution of a methyl group at C_5 the largest changes are in the internal angle at C_4 and C_5 which alter by ~1°. Bond lengths change by less than 0.01 Å. Substitution of fluorine at C_5 changes the internal angles at C_4 and C_5 by ~1° while all other changes are less than 1° or 0.01 Å.

Relative Tautomeric Stabilities. The calculated total energies of the tautomers studied are given in Table VI. Their calculated relative stabilities are compared with previous theoretical results and available experimental data in Tables VII–X. Tables XI and XII give the calculated dipole moments and equilibrium constants involving the most stable tautomer of each species. The latter values have been estimated simply from the calculated energies of Table VI. We thus neglect zero point energy differences, which for the lactam/lactim type equilibria discussed here have been estimated to be ~3 kJ mol^{-1},[26] the temperature dependence of the enthalpy, which has been shown to be a reasonable assumption,[26] and the $T\Delta S$ term for the tautomeric equilibria. For U1 ⇌ U3, this final term has been estimated to

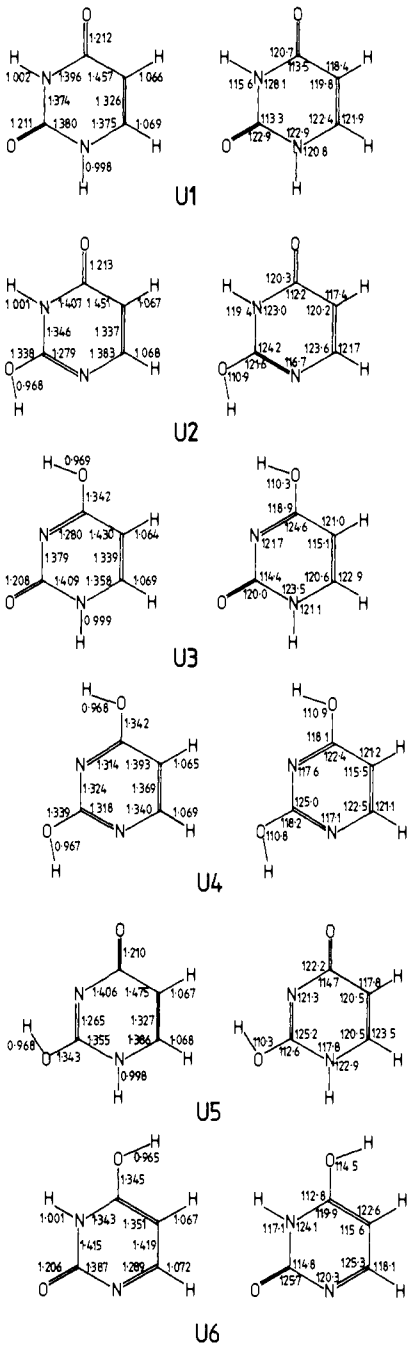


Figure 1. Optimized bond lengths and angles for uracil tautomers U1, U2, U3, U4, U5, and U6.

be ~1.3 kJ mol^{-1}.[42] Thus, the enthalpic term is dominant in the determination of the equilibrium constant.

The major conclusions from our calculated tautomer energetics are that (i) the diketo tautomers of uracil, thymine, and 5-fluorouracil (U1, T1, and FU1) are the most stable, in agreement with experiment,[3] (ii) tautomers U2, T2, and FU2 are the most stable rare tautomers in the gas phase and not the more commonly studied tautomers U3, T3, and FU3, (iii) the order of stability of the six tautomers of uracil is unaffected by substitution of either a methyl group or fluorine at the 5-position, (iv) the relative stability of equivalent tautomers of uracil, thymine, and 5-fluorouracil differs by less than 8 kJ mol^{-1} except for the tautomers U6 and T6 where the difference is ~17 kJ mol^{-1}, and (v) at least two and possibly three tautomers of cytosine should be present in the gas phase at normal temperatures.

(42) Shibata, M.; Zielinski, T. J.; and Rein, R. *Int. J. Quantum Chem.* **1980**, *18*, 323.

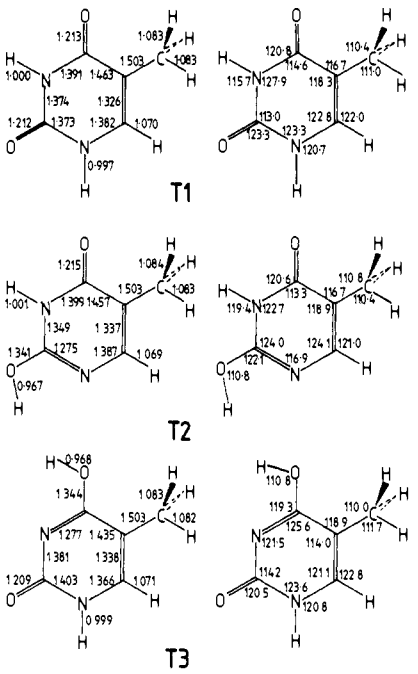


Figure 2. Optimized bond lengths and angles for thymine tautomers T1, T2, and T3. The C_4–C_5–C_9–H torsion angles for T1, T2, and T3 are 59.4, 59.7, and 59.0°, respectively.

Table X. Relative Energies (kJ mol^{-1}) of Cytosine Tautomers

	tautomer					
method	C1	C2	C3	C4	C5	C6
CNDO/2[a 12]	0.0	111.0				
CNDO/2[a 13]	0.0	1.6				
CNDO/2[a 15]	0.0	80.6	6.2	61.3		
Hückel[a 11]	0.0	9.2	19.7	3.1		
MINDO/2[17]	0.0	43.1				
MINDO/3[18]	0.0	19.7	37.2	15.5	80.3	83.7
MNDO[24]	0.0	20.1	−41.6	13.8	29.8	43.2
STO-3G[a 16]	0.0	46.5	−74.2	25.0	48.4	99.1
ab initio[22]	0.0	−0.6				
MNDO// 3-21G[b]	0.0	0.5	41.7 (B)	49.7	87.5	96.5
3-21G[b]	0.0	1.7	15.9 (A) 19.3 (B)	29.8	76.2	94.8

[a] Optimized geometries have not been used. [b] This work.

Table XI. Calculated Dipole Moments (D) for the Tautomers of Uracil, Thymine, 5-Fluorouracil, and Cytosine

	tautomer					
	1	2	3	4	5	6
uracil	4.8	3.4	5.2	1.6	6.5	7.6
thymine	4.7	2.9	5.4	1.8	6.2	7.8
5-fluorouracil	4.3	4.7	3.6	1.3	7.1	5.8
cytosine	7.2	5.3	3.7 (A) 5.2 (B)	8.6	1.8	4.5

Uracil, Thymine, and 5-Fluorouracil. The tautomerism of uracil has been the subject of several previous theoretical investigations[11,12,14,17-21,23,24] (Table VII), which have concentrated on the equilibrium U1 ⇌ U3 since only those tautomers protonated at N_1 (U1, U3, U5) can occur in the nucleic acids and U3 was assumed to be more stable than U5. Our work shows U3 to be more stable than U5 by 28 kJ mol^{-1} and less stable than U1 by 81.8 kJ mol^{-1} in the gas phase. The latter figure is 50 kJ mol^{-1} greater than that predicted by the STO-3G results.[21] However,

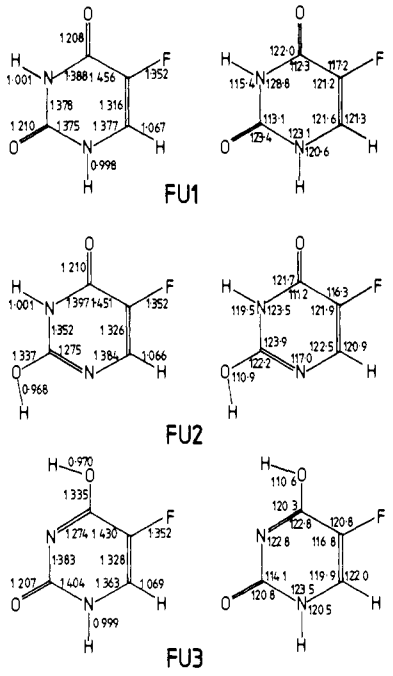


Figure 3. Optimized bond lengths and angles for 5-fluorouracil tautomers FU1, FU2, and FU3.

Table XII. Predicted Equilibrium Constants with Respect to the Most Stable Tautomer (T = 298 K)

	tautomers				
	2	3	4	5	6
uracil	2×10^{-13}	5×10^{-15}	2×10^{-18}	7×10^{-20}	1×10^{-21}
thymine	4×10^{-13}	6×10^{-16}	3×10^{-19}	2×10^{-19}	1×10^{-24}
5-fluoro-uracil	2×10^{-12}	2×10^{-15}	9×10^{-18}	2×10^{-18}	8×10^{-21}
cytosine	0.5	2×10^{-3} 4×10^{-4}	6×10^{-6}	4×10^{-14}	2×10^{-17}

the STO-3G basis has previously been shown to overestimate the stability of the lactim tautomer by 50 kJ mol^{-1} for the 2-pyridone/2-hydroxypyridine equilibrium.[28]

Beak et al.[10] have estimated ΔH for U1 $\rightleftharpoons$ U3 and U1 $\rightleftharpoons$ U4 as 79.5 (±25.1) and 92.0 (±41.8) kJ mol^{-1}, respectively, in good agreement with our corresponding results of 81.8 and 101.1 kJ mol^{-1}. The error found for corresponding calculations of the 2-pyridone/2-hydroxypyridine system, 10 kJ mol^{-1},[29] suggests that the uncertainty associated with our calculated values is less than that associated with the experimental values.

In agreement with the results presented here the rare lactim tautomers of uracil have not been directly observed in either the gas phase[6,7] or solution.[3] This has led to the conclusion that the tautomeric equilibria are unaffected by phase change.[6,7] However, comparison of the equilibrium constants estimated in the gas phase for the equilibria U1 $\rightleftharpoons$ U2 and U1 $\rightleftharpoons$ U3, 2×10^{-13} and 5×10^{-15}, respectively, with the experimental equilibrium constants in solution[43] of 8.7×10^{-6} and 2.5×10^{-4} indicates that on going from the gas phase to solution these equilibrium constants change by factors of 10^7 and 10^{11}, respectively. It is also apparent that the order of stability of tautomers 2 and 3 is reversed on going from the gas phase to solution.

The observation that U2, T2, and FU2 are more stable than U3, T3, and FU3, respectively, implies that the proton at N_1 is more acidic than that at N_3. This is in qualitative agreement with the results of Del Bene, who has shown that H bonding of water to $H(N_1)$ is more stable than that to $H(N_3)$ for the most stable tautomers of uracil,[44] thymine,[45] and 5-fluorouracil.[46]

(43) Poulter, C. D.; Frederick, G. D. *Tetrahedron Lett.* **1975**, *26*, 2171.
(44) Del Bene, J. E. *J. Comput. Chem.* **1981**, *2*, 188, 416.
(45) Del Bene, J. E. *J. Chem. Phys.* **1982**, *76*, 1058.

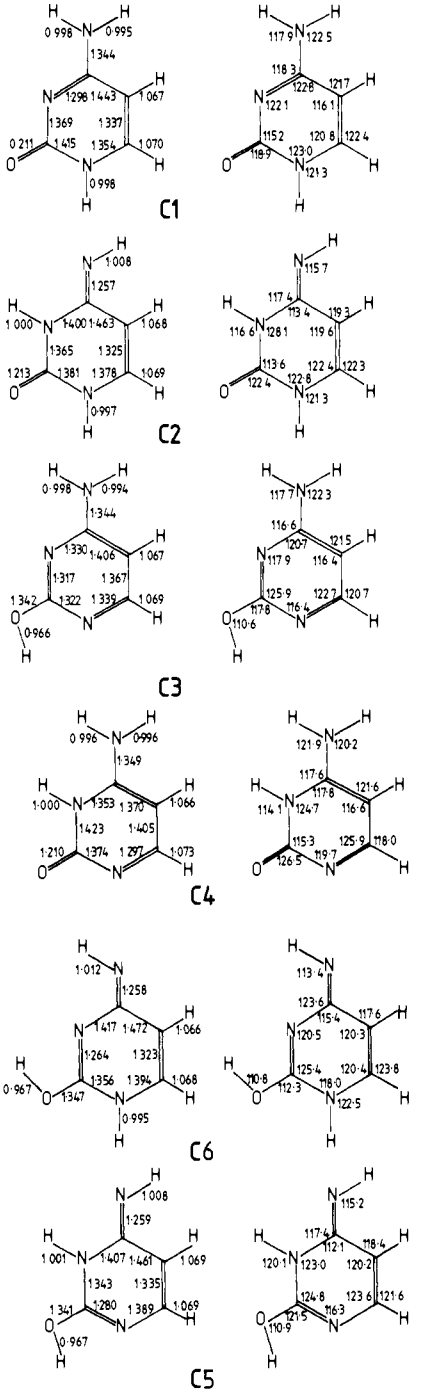


Figure 4. Optimized bond lengths and angles for cytosine tautomers C1, C2, C3, C4, C5, and C6.

Substitution of either a methyl group or a fluorine atom at C_5 appears to decrease the stability of the 4-lactim tautomer of thymine and 5-fluorouracil (T3 and FU3) relative to U3 while it increases the stability of the 2-lactim tautomers, T2 and FU2, relative to U2. These changes in relative stability are small and only qualitative conclusions can be drawn from them, the most significant being that in the gas phase FU3 is not more stable relative to FU1 than is U3 compared to U1. The *increased* stability of FU3 relative to U3 has been postulated as an explanation for the mutagenic effect of 5-fluorouracil.[31]

The possibility of a favorable intramolecular interaction between the electronegative fluorine and the hydroxyl group at C_4 in FU3 has been considered by partial optimization of this tautomer with O–H trans to N_3–C_4. The relative stability of this conformation

(46) Del Bene, J. E. *J. Phys. Chem.* **1982**, *86*, 1341.

Table XIII. Base–Water Interaction Energy (kJ mol^{-1}) from the Reaction Field Continuum Model

	tautomer					
	1	2	3	4	5	6
uracil	−25.7	−13.0	−30.4	−2.9	−47.5	−65.0
thymine	−26.4	−10.1	−35.2	−3.9	−46.4	−73.4
5-fluorouracil	−20.6	−24.7	−14.5	−1.9	−56.5	−37.7
cytosine	−59.2	−32.1	−15.7 (A)	−84.5	−3.7	−23.3
			−31.1 (B)			

Table XIV. Relative Tautomeric Stabilities (kJ mol^{-1}) Corrected for Base–Water Interaction

	tautomer					
	1	2	3	4	5	6
uracil	0.0	84.8	77.1	123.9	87.5	80.3
thymine	0.0	87.3	77.9	128.3	87.2	89.5
5-fluorouracil	0.0	62.1	89.9	116.0	65.6	97.7
cytosine	0.0	28.8	59.4 (A)	4.5	131.7	130.7
			47.4 (B)			

is 109.4 kJ mol^{-1}, 26 kJ mol^{-1} less stable than the cis conformation, ruling out the possibility of such an interaction.

Cytosine. In aqueous solution C1 is found to be the dominant tautomer, by factors of about 10^3 and $10^{4.6}$ with respect to C4 and C2.[3] Tautomers C3, C5, and C6 are not present in detectable amounts. Dreyfus et al.[47] have estimated the equilibrium constant $K_{C4/C1}$ to be 2.5×10^{-3} in aqueous solution and the associated ΔH to be 13.0 (±0.4) kJ mol^{-1}. Their studies of 3-methylcytosine show that for the equilibrium C4 ⇌ C2, C4 is preferred in aqueous solution ($K_{C2/C4} = 3 \times 10^{-2}$) while in nonpolar solvent C2 is the dominant tautomer. $K_{C2/C1}$ is thus estimated to be 10^{-4}–10^{-5} in aqueous solution.

From these data we may infer that in aqueous solution the order of stability is C1 > C4 > C2 while C3, C5, and C6 are not detected. However, we predict the order of stability in the gas phase to be C1 > C2 > C3 > C4. Furthermore, comparison of the gas-phase equilibrium constants, $K_{C2/C1} = 0.5$, $K_{C3/C1} = 2 \times 10^{-3}$, $K_{C4/C1} = 6 \times 10^{-6}$, and $K_{C2/C4} = 8 \times 10^4$, with the experimental values given above implies that large changes may occur on going from the gas phase to solution. In view of this, we now make theoretical estimates of such changes.

Solvent Effects. Solvent effects have previously been considered to comprise two major components,[10] namely electrostatic solvent–solute interactions and hydrogen bonding. The hydrogen-bonding effects cannot be estimated in a quantitative manner, without further large-scale calculations. The electrostatic solvent–solute effects, however, are readily estimated by the reaction field continuum model,[48] using the dipole moments reported herein (Table XI) and molecular polarizabilities calculated by the method of Miller and Savchik.[49] In Table XIII the values of the base–water interaction for each tautomer thus calculated are presented, using a value of 78.54 for the dielectric constant of water and taking a spherical cavity radius of 3.3 Å. These values allow the relative stabilities of the tautomers in aqueous solution to be estimated as given in Table XIV. Consideration of the solvent causes some reordering of the stability of the uracil tautomers, which is now predicted to be U1 > U3 > U6 > U2 > U5 > U4, in agreement with the experimental results.[42,50] There is still a large discrepancy between our predicted value for $K_{U3/U1}$ in solution (3×10^{-14}) and that measured experimentally (2.5×10^{-4}) which may arise from hydrogen bonding effects. It should be noted that for uracil the solvent effects must favor the lactim tautomers (U2 and U3) relative to the lactam tautomer (U1), in contrast to the situation for the 2-pyridone/2-hydroxypyridine equilibrium where solvation clearly favors the lactam tautomer.[5]

The stabilization of the thymine tautomers predicted upon solvation is similar to that predicted for uracil. In both cases the 4-lactim tautomers (U3, T3) are predicted to be the most stable rare tautomer in solution. In the case of 5-fluorouracil, however, the 2-lactim tautomers FU2 and FU5 are both predicted to be considerably more stable than FU3. Both FU3 and FU5 may exist within nucleic acids, and it is apparent that depending upon the polarity of the environment either FU3 or FU5 will be the more stable. Previous discussions of the role of lactim tautomers of 5-fluorouracil in mutation have been limited to consideration of FU3.[14,21,31]

For cytosine, estimation of the solvent effects in this manner provides excellent agreement with the experimental data. Thus, the order of stability is now predicted to be C1 > C4 > C2 and C3 should not be observed. We predict the values of $K_{C4/C1}$, $K_{C2/C4}$, $K_{C2/C1}$, $K_{C3/C1}$ to be 0.16, 5×10^{-5}, 9×10^{-6}, and 4×10^{-11}, respectively, in aqueous solution. These values are within a factor of $\sim 10^2$ of the experimental values, corresponding to an error in ΔG of the order of 10 kJ mol^{-1}.

Thus application of the continuum reaction field model leads to an explanation of the change in the order of tautomeric stability on going from gas phase to solution. However, such a treatment lacks explicit consideration of base–water hydrogen-bonding effects,[51] and the associated entropy changes so that relative stabilities predicted in water are considerably less reliable than those predicted in the gas phase.

Biological Significance. From the equilibrium constants that have been estimated in Table XII it is predicted that the concentrations of the rare enol tautomers U3, T3, and FU3 are $\sim 10^3$ times smaller than the observed frequency of spontaneous mutation (10^{-8}–10^{-11}). However, the concentration of the cytosine tautomer C2 should be $\sim 10^8$ times greater than the observed frequency of mutation. These discrepancies are far greater than any that might be caused by errors in the estimation of the gas-phase equilibrium constants and indicate clearly that other factors apart from the relative stability of the tautomers of the pyrimidine bases themselves are of fundamental importance in determining the stability of the individual tautomers within the nucleic acids. If the relative stability of the tautomers themselves is all that is considered approximately one-third of the cytosine bases in DNA might exist as the imine tautomer C2 and therefore be expected to mispair during replication.

The factors most likely to influence the tautomeric equilibria are substitution at the N_1 position, hydrogen bonding, association, base pairing, solvent polarity, and other intramolecular interactions within the nucleic acids.

The effect of substitution at the N_1 position is largely unknown although bases substituted at N_1 with a methyl group or a sugar are often used in experimental studies of tautomerism.[3] Association has been shown to be important in the consideration of the 2-pyridone/2-hydroxypyridine system and to favor the lactam tautomer.[5] The tautomeric equilibria discussed here have been shown to be very sensitive to phase change and solvent polarity.[47] With use of the gas-phase equilibrium constants presented here and a simple model of solvation some attempt has been made to quantify the effects associated with phase change.

A number of other theoretical studies employing a variety of methods have previously been concerned with solvation of the pyrimidine bases[41,44-46,52] and with base pairing.[53] However, these studies have been limited to consideration of the most stable tautomers (U1, T1, FU1, and C1). Application of theoretical methods to the consideration of solvation and base pairing of the

(47) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J. Am. Chem. Soc.* **1976**, *98*, 6338.
(48) Rein, R.; Renugopalakrishnan, V.; Nir, S.; Swissler, T. J. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1975**, *2*, 99.
(49) Miller, K. J.; Savchik, J. A. *J. Am. Chem. Soc.* **1979**, *101*, 7206.
(50) Sukhorukov, B. I.; Gukovskaya, A. S.; Nekrasova, V. G.; Antonovskii, V. L.; *Biophysics (Biofizika)* **1974**, *19*, 808.
(51) Richards, W. G. "Quantum Pharmacology"; Butterworths & Co.: London, 1983.
(52) Scordamaglia, R.; Cavallone, F.; Clementi, E. *J. Am. Chem. Soc.* **1977**, *99*, 5545. Pullman, A.; Perahia, D. *Theor. Chim. Acta* **1978**, *48*, 29. Pullman, B.; Miertus, S.; Perahia, D. *Ibid.* **1979**, *50*, 317. Clementi, E.; Corongiu, G. *J. Chem. Phys.* **1980**, *72*, 3979.
(53) Clementi, E.; Mehl, J.; von Niessen, W. *J. Chem. Phys.* **1971**, *54*, 508.

rare tautomers is likely to provide further useful information concerning the significance of tautomerism as a mechanism of mutation in nucleic acids.

Influence of Molecular Geometry on Tautomer Energetics. The use of molecular geometries determined in a minimal basis or by semiempirical methods followed by single-point energy evaluation in an extended basis has obvious computational advantages for the calculation of tautomer energetics. In Table X we show such results for C1-C6 using a MNDO optimized geometry followed by single-point 3-21G basis calculations (MNDO//3-21G). Similar results are shown in Table VII for U1 and U3 (STO-3G//3-21G). For cytosine, the largest errors (for C3 and C4) in the relative energies are $\sim$20 kJ mol^{-1} compared to our results using geometries optimized at the 3-21G level, although the order of stability of the tautomers is not altered. For U1 and U3 the STO-3G//3-21G calculations give an error of 11 kJ mol^{-1}, rather larger than the value of 1.2 kJ mol^{-1} found for the 2-pyridone/2-hydroxypyridine system.[28] These errors are considerably smaller than those resulting from the use of MNDO or STO-3G wave functions alone, both methods overestimating the stability of the lactim tautomer.

Basis Set Extension and Correlation Effects. We have not here considered larger basis sets than the 3-21G, nor the role of correlation effects. Geometry optimization of pyrimidine has been carried out by using a larger basis than that used here.[54] It was found that to obtain a comparable accuracy for the bond angles at nitrogen to that obtained for the angles at carbon, it was necessary to include polarization functions on the nitrogen atom. A similar effect is suggested from a comparison of the optimized structures obtained herein with the experimental geometries where the predicted CNC angles are too large by about 2°. Previous studies of 2-pyridone/2-hydroxypyridine and the corresponding 4-substituted isomers[28,29] suggest that the addition of polarization functions consistently favors the lactim tautomer by $\sim$10 kJ mol^{-1}. This is of particular significance for tautomer C3, for if it is stabilized relative to C1 by an additional 10 kJ mol^{-1} we would expect to observe it in the gas phase. However, calculations using Moller-Plesset perturbation theory[28] suggest that correlation effects favor the lactam tautomer by $\sim$5 kJ mol^{-1}.

(54) Pang, F.; Pulay, P.; Boggs, J. E. *J. Mol. Struct. (Theochem.)* **1982**, *88*, 79.

Conclusions

The major conclusions of the present theoretical study are the following: (1) Optimization at the 3-21G level yields molecular geometries in good agreement with average crystallographic values for uracil and cytosine, and superior to those previously obtained theoretically. (2) The most stable tautomers are predicted to be U1, T1, FU1, and C1 in agreement with experiment. The calculated relative energies of U1, U3, and U4 are in excellent agreement with experiment. (3) Substitution of uracil at the 5-position by CH_3 or F does not change the order of the stabilities of the tautomers. (4) The tautomeric equilibria of both uracil and cytosine are sensitive to phase change. (5) Use of the reaction field continuum model successfully explains the reordering of the relative tautomeric stabilities on passing from the gas phase to solution.

Registry No. Uracil, 66-22-8; cytosine, 71-30-7; thymine, 65-71-4; 5-fluorouracil, 51-21-8.

Influence of d-Orbital Occupancy on the Geometry of Pentacoordinated Molecules[1]

Robert R. Holmes

Contribution from the Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003. *Received July 13, 1983*

Abstract: A quantitative assessment of solid-state structural data on five-coordinated compounds of both transition and main group elements establishes the geometrical change in the square pyramid as a function of d-orbital configuration. In terms of the trans basal angle θ, a variation in the range 140-174° is obtained. These angle variations are interpreted in terms of nonbonded repulsions between d-orbital electron density and bond electron density. Structural distortion from the trigonal-bipyramidal and square-pyramidal geometries are determined for the five-coordinated compounds by using a dihedral angle method. It is found that the local distortion coordinate for transition-metal complexes approximates the Berry intramolecular exchange coordinate. Main group pentacoordinated compounds follow this coordinate more closely.

It is known[2] that the square-pyramidal geometry assumed by some pentacoordinated phosphorus compounds has the phosphorus atom located out of the basal plane such that the trans basal angles, θ_{15} and θ_{24}, average 152°. This value was predicted by Zemann[3] on the basis of a point-charge model. X-ray investigations on a variety of five-coordinated phosphorus compounds further show that the structural form assumed varies from near the idealized trigonal bipyramid (TBP) to the square pyramid (SP).[2,4] These structures lie along a C_{2v} coordinate connecting the two limiting geometries and are supportive of the Berry pseudorotation mechanism[5] frequently invoked[6] to interpret dynamic NMR data indicative of intramolecular ligand exchange.

Although a great deal of structural work has been reported for five-coordinated transition-metal compounds, no quantitative assessment of the "preferred" geometry of the square-pyramidal form is available. The literature contains isolated references of possible variations. For example, high- and low-spin square-pyramidal Co(II) and Ni(II) complexes have been classified by Orioli[7] in terms of metal atom displacement from the basal plane.

(1) Pentacoordinated Molecules. 55. Previous paper in the series: Deiters, J. A.; Holmes, R. R. *J. Am. Chem. Soc.* **1984**, *106*, 3307-3309.
(2) Holmes, R. R. *ACS Monogr.* **1980**, *175*, 44.
(3) Zemann, J. *Z. Anorg. Allg. Chem.* **1963**, *324*, 241.
(4) Holmes, R. R.; Deiters, J. A. *J. Am. Chem. Soc.* **1977**, *99*, 3318.
(5) Berry, R. S. *J. Chem. Phys.* **1960**, *32*, 933.
(6) See ref 2, Chapter 3.
(7) Orioli, P. L. *Coord. Chem. Rev.* **1971**, *6*, 285.

0002-7863/84/1506-3745$01.50/0