Quantum Chemical Investigation of the Structure and Stability of All Geometric Isomers and Conformers of All Tautomeric Forms of Thymine

Tae-Kyu Ha and H. H. Gunthard*

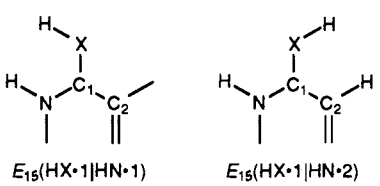
Contribution from the Physical Chemistry Laboratory, ETH Zentrum, CH-8092 Zuerich, Switzerland

*Received March 22, 1993**

Abstract: Optimized structures and electronic energies are reported for all geometric isomers of all five tautomeric forms of thymine based on quantum chemical HF/6-31G** and (pointwise) MP2/HF/6-31G** computations. Furthermore, electric dipole moments of all isomers and rotational constants and electric field gradients at the ring nitrogen atoms of the four most stable isomers are included. Similarly, electronic energies of all geometric isomers of two more tautomers of both uracil and cytosine are given, thus completing a previously published set of quantum chemical data for these two nucleic acid bases to comprehend conversion energies of geometric isomers of all tautomers. A consistent system of contribution terms for all three bases is determined, from which conversion energies of geometric isomers may be additively expressed within error limits (≈0.5 kcal/mol). The contribution terms represent either repulsive interactions between hydrogen atoms bound to ring atoms and to OH or NH substituents or attractive interactions ("intramolecular H-bonds") between hydroxy or imino hydrogen atoms and lone-pair electrons localized at adjacent N atoms. Alterations of internal structural parameters accompanying anti–syn conversion of geometric isomers of thymine are described by linear regression expressions and mechanically interpreted in terms of repulsive and attractive interactions. Also, electric field gradients at the N atoms are shown to be correlated closely to these quantities. Predicted electric dipole moments and rotation constants are found to approximate closely empirical data (where available). However, predicted internal structural parameters were found to deviate significantly from X-ray data for thymine.

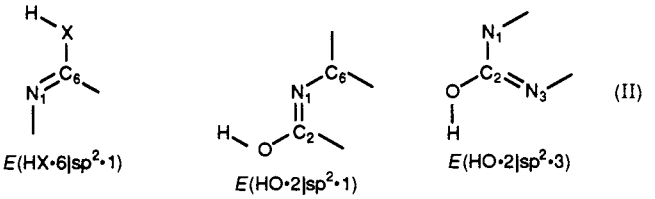
Introduction

In a recent paper, a quantum chemical study of nine geometric isomers of both cytosine and uracil at the HF/6-31G** (MP2/HF/6-31G**) level was reported.[1] The main results were the following. (i) The establishment of correlations between geometric isomerism (conformations) of NH and OH substituents of cytosine and uracil and internal structural parameters (ISPs, bond lengths and bond angles). Such correlations (structural relaxations) were most clearly expressed for ISPs determining local nuclear configurations in the regions of the substituents. (ii) The determination of a small set of contributions (increments), from which the electronic interconversion energies of geometric isomers may be additively reproduced. Such contributions were found to be directly perceptible from the structures of the geometric isomers, either as repulsive interactions between H atoms of NH or OH substituents and H atoms bound to adjacent N or C atoms of the ring or as attractive interactions between H atoms of such substituents and lone-pair electrons localized at adjacent N atoms of the ring. Repulsive interactions are associated with structural fragmemts (X = N:, O:) shown below:



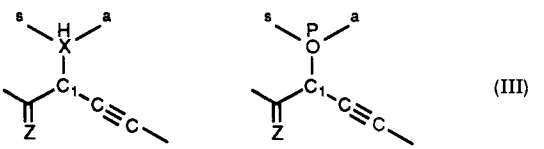
E_{15}(HX•1|HN•1) E_{15}(HX•1|HN•2) (I)

Likewise, attractive interactions are associated with structural schemes of the type shown below:



E(HX•6|sp²•1) E(HO•2|sp²•1) E(HO•2|sp²•3) (II)

The search for additive contributions to conversion energies emerged in two earlier quantum chemical studies of the conversion energies of geometric isomers of imines of simple acyclic aldehydes and ketones and of analogous isoelectronic protonated carbonyl compounds.[2] The conversion energies of geometric isomers of these compounds were found to fall into a few classes with respect to (wrt) magnitude and sign, analogous (isoelectronic) imino and O-protonated compounds possessing quite similar conversion energies. These observations suggested the existence of increments from which the conversion energies of pairs of geometric isomers may be rebuilt. However, determination of such contributions was hindered by the fact that the number of equations expressing conversion energies by linear combinations of even a few contributions is smaller than the number of the latter. To arrive at a unique solution, the syn → anti conversion energies of a number of model compounds of the type (X = O, N; Z = CH_2, O, N; p = proton; s and a denote syn and anti positions) shown in (III) were computed quantum chemically at the same level of



(III)

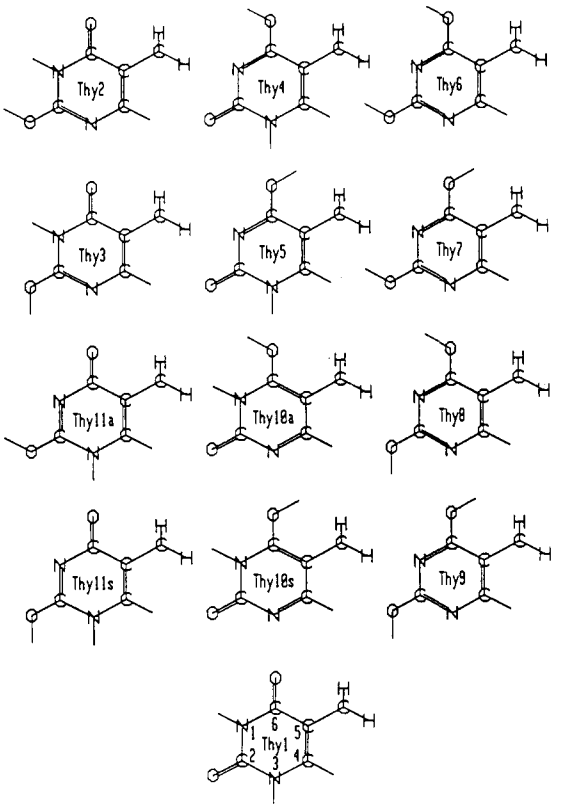
approximation.[3] For the anti isomer of these models, repulsion (and attraction) between H (of the HX group) and the acetylenic

* Abstract published in *Advance ACS Abstracts*, October 15, 1993.

(1) Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct. (THEOCHEM)* **1992**, *276*, 209–249.

0002-7863/93/1515-11939$04.00/0

Chart I. Thymine Isomers Thy1–Thy11[a]



[a] CH_3 group eclipsing C4=C5.

chain was postulated to be vanishingly small. By this procedure, zeroth-order estimates for repulsive interactions of the imino and O-protonated analogs were provided. Estimation by least-squares or maximum likelihood techniques yielded a set of increments, from which computed isomerization energies may be reproduced within ≤0.5 kcal/mol. For the imino and O-protonated compounds, only repulsive terms occur, which, within the approximations HF/6-31G*, HF/6-31G**, and MP2/HF/6-31G**, seemed fairly independent of basis sets and applicable rather generally. This encouraged the use of these repulsions as zeroth-order estimates for repulsive interactions occurring with pyrimidine nucleic acid bases.

Using a similar approach, a set of additive contributions to the conversion energies of all geometric isomers of three cytosine and three uracil tautomers was derived. Analogous increments for cytosine and uracil were found to be equal within the error limits of the linear model. Besides repulsive contributions, it proved essential to introduce *attractive* interactions of the type mentioned above. The latter may be classified as analogs of intramolecular H-bonds with strongly bent geometries and bonding energies in the range 3–5 kcal/mol.

The present work extends the earlier investigations in three ways. (i) By providing optimized structural and electronic energy data for all five thymine tautomeric forms, including all their geometric isomers and two methyl conformations. Furthermore, the electronic energies of all geometric isomers of the last two tautomers of cytosine and uracil, not treated in the foregoing paper, are reported. (ii) By completing the structure–geometric isomerism correlations (structural relaxation upon conversion of isomers) including the conformers of the methyl group. (iii) By determining of a linear contribution model of the conversion energies of geometric isomers of thymine and of all diastereomers

Chart II. Cytosine Isomers C10 and C11

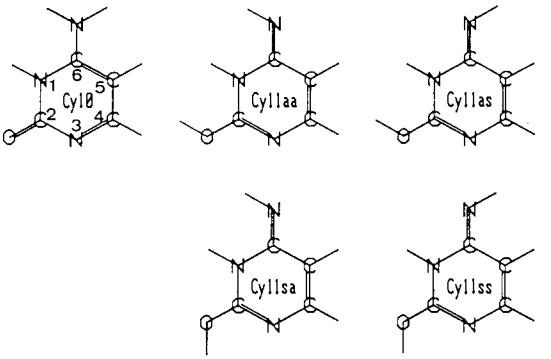
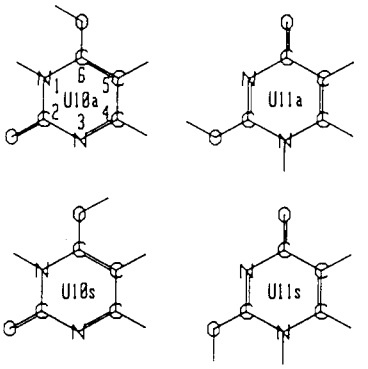


Chart III. Uracil Isomers U10 and U11



of all tautomers of cytosine and uracil. Thus a consistent set of increments of isomerization energies of geometric isomers of all three pyrimidine nucleic acid bases is provided.

The thymine compounds treated in this work are shown in Chart I and the cytosine and uracil isomers in Charts II and III, respectively.

The pyrimidine nucleic acid bases have been subject to numerous quantum chemical investigations. For a rather comprehensive citation of earlier work, the reader is referred to the *Quantum Chemistry Literature Data Base*.[4,5] Papers of direct relevance to the present work on thymine were published by several workers, e.g., Scanlan et al.,[6] Foerner et al.,[7] Eisenstein,[8] Urano et al.,[9] Basch et al.,[10] and Brown.[11] To our knowledge, no investigation of all diastereomers and conformers of all thymine tautomers at the presently highest feasible quantum chemical approximation has been published hitherto. The present paper should fill this gap. Though some thymine tautomers have high relative electronic energies and, therefore, in thermal equilibrium will be present in low concentration, it might be possible to populate unstable tautomers and their geometrically isomeric variants by UV and IR irradiation, thereby making these species amenable to experimental investigation.[12-14]

Computational Details

Ab Initio Calculations. Equilibrium molecular structures and electronic energies of all geometric isomers of the five tautomeric forms of thymine (5-methyluracil), including minimum and maximum energy conformations

(2) Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct.* (*THEOCHEM*) **1992**, *259*, 229–256 (Part I).
(3) Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct.* (*THEOCHEM*) **1992**, *276*, 187–208 (Part II).
(4) *Quantum Chemistry Literature Data Base, Bibliography of ab initio calculations 1978–1980*; Elsevier: Amsterdam, 1982. Supplements to *J. Mol. Struct.* (*THEOCHEM*): **1982**, *91*; **1983**, *106*; **1984**, *119*; **1985**, *134*; **1986**, *148*; **1987**, *154*; **1988**, *182*; **1989**, *203*; **1990**, *211*; **1991**, *252*.
(5) For references to work relating to cytosine and uracil, cf. ref 1.
(6) Scanlan, M. J.; Hillier, I. H. *J. Am. Chem. Soc.* **1984**, *101*, 3737.
(7) Foerner, W.; Ladik, J.; Otto, P.; Cizek, J. *Chem. Phys.* **1986**, *97*, 251.
(8) Eisenstein, M. *Int. J. Quantum Chem.* **1988**, *33*, 127.
(9) Urano, S.; Yang, X.; LeBreton, P. R. *J. Mol. Struct.* **1989**, *214*, 315.
(10) Basch, H.; Garner, D. R.; Jasien, P. G. *Chem. Phys. Lett.* **1989**, *163*, 514.
(11) Brown, R. D. *Chem. Soc. Commun.* **1989**, 37.
(12) Frei, H.; Pimentel, G. C. *Annu. Rev. Phys. Chem.* **1985**, *36*, 491.
(13) Raesanen, M.; Kunttu, H.; Murto, J. *Laser Chem.* **1985**, *9*, 123.
(14) Hollenstein, H.; Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct.* **1986**, *146*, 289.

of the methyl group, were determined by ab initio methods. This comprises both lactam–lactime and imine–amine tautomerism and anti–syn isomerism of OH groups, cf. Chart I. The molecular geometries were optimized with respect to all internal structural parameters (ISPs) by the force method with analytical gradient. The 6-31G** basis set[15] was employed for geometry optimization. To study the influence of electron correlation on calculated relative energies, single-point computations were performed at the MP2/6-31G** level of theory using HF/6-31G** optimized geometries. Characterization of stationary points and computation of harmonic frequencies and zero-point vibrational energies (ZPE) were made at the HF/6-31G** level of theory. The GAUSSIAN 90 program package[16] was used for the ab initio calculations.

Furthermore, optimized structures and electronic energies of the five cytosine isomers and four uracil isomers pictured in Charts II and III, respectively, were computed; the latter data complete the results reported earlier[1] to comprehend now all cytosine and uracil tautomers and their diastereomers.

Analysis of Quantum Chemical Data. To express correlations between any ISP ξ on one hand and geometric isomerism of OH substituents in ring positions 2 and 6 and the conformation of the CH_3 group in the 5 position on the other hand, linear relationships of the form

$$\xi(x_2,x_6,x_5) - \xi(0,0,0) = a_2x_2 + a_6x_6 + a_5x_5 \qquad (1)$$

will be used, where (cf. Figure 1 for notation of ring positions) $x_2 = 0$ (1) for the anti (syn) position of the 2-OH group, $x_6 = 0$ (1) for the anti (syn) position of the 6-OH group, and $x_5 = 0$ (1) for CH_3 eclipsing (staggering) $C_4{=}C_5$. For isomers with C:O groups in the 2- or 6-position, the corresponding terms in (1) will be cancelled. Determination of the coefficients will be made by least-squares methods; positive (negative) coefficients express direct (anti) correlation.

Determination of additive contributions to conversion energies of geometric isomers related to OH substituents in the 2- and 6-positions will follow the procedure outlined in detail in ref 1 but further will include 1,6-H,H(E_{16}) repulsive interactions between HO·6 and H-atoms of the CH_3·5 group, depending on conformation. Very briefly, the assumptions of the additive model are recapitulated. (i) The total electronic energy of every isomer $\mathcal{S}_m$ in a set ($\mathcal{S}$) of geometric isomers associated with a tautomer ($\mathcal{S}$) is built up from a term ($E_{\mathcal{S}}$) characteristic of the set and a few contributions ($E_{\mathcal{S}_m,ik}$) characteristic of the isomer $\mathcal{S}_m$. (ii) Each contribution (increment) $E_{\mathcal{S}_m,ik}$ is either a repulsive interaction (>0) between H atoms of OH substituents and H atoms bound to adjacent ring atoms or H atoms of the CH_3 group or an attractive interaction (<0) between H atoms of OH substituents and lone-pair electrons centered at adjacent ring N atoms. (iii) Any conversion energy of geometric isomers in set $\mathcal{S}$ is then expressed by a linear equation of the form

$$\Delta\mathcal{E} = -(E_{\mathcal{S}} + \sum_{ik} E_{\mathcal{S}_m,ik})_{\text{educt}} + (E_{\mathcal{S}} + \sum_{ik} E_{\mathcal{S}_m,ik})_{\text{product}}$$

$$= -(\sum_{ik} E_{\mathcal{S}_m,ik})_{\text{educt}} + (\sum_{ik} E_{\mathcal{S}_m,ik})_{\text{product}} \qquad (2)$$

Typical nuclear configurations, to which repulsive and attractive interactions are attributed, are sketched in Charts I and II shown in the Introduction. Estimation of numerical values of contributions is made by least-squares or maximum likelihood techniques. Zeroth-order estimates of specific interactions are taken from appropriately chosen model compounds. The majority of model compounds used were reported earlier. In the case of thymine, a few further model compounds are required; these are described in the Appendix. It should be clear that the use of model compounds for generation of zeroth-order estimates of interaction terms introduces a bias in the determination of the latter. This procedure should be tolerable, since it permits determination of a consistent system of interactions holding for all pyrimidine derivatives *and* for the simple acyclic imines and O-protonated carbonyl compounds treated so far.

Results

Most internal structural parameters (ISPs) of the optimized structures of the 13 isomeric thymines with the methyl group eclipsing the double bond $C_4{=}C_5$ (ec conformer) are shown in

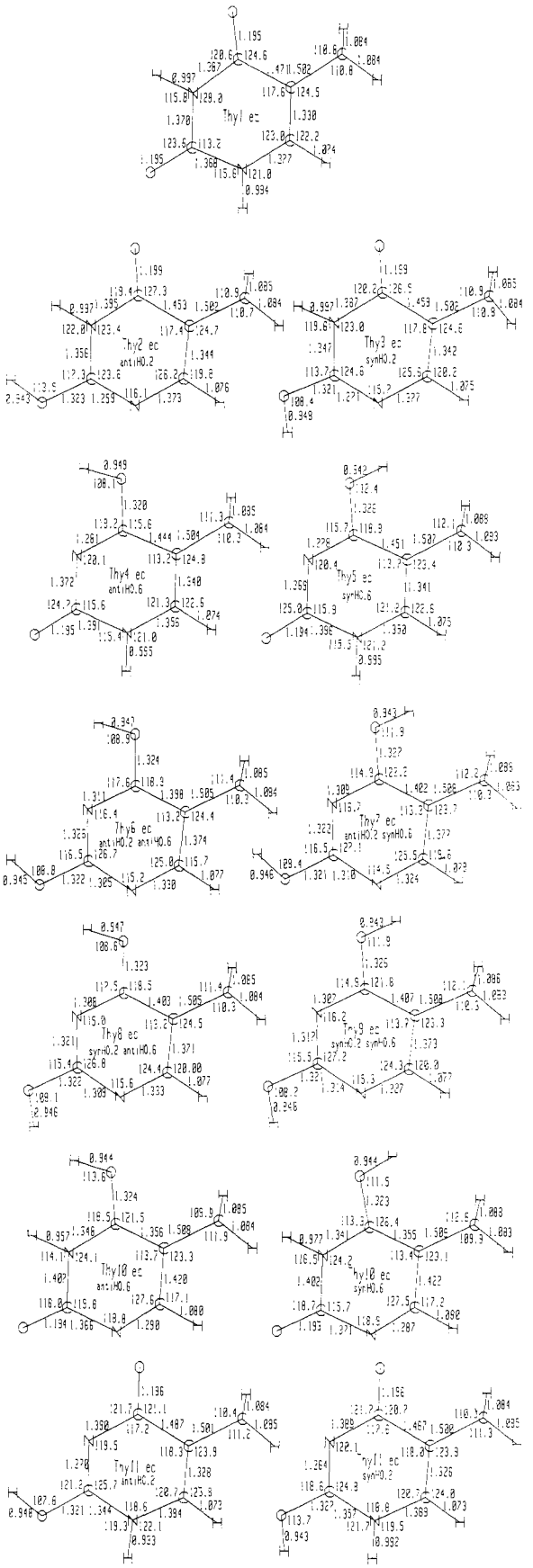


Figure 1. Optimized internal structural parameters (ISPs) of thymine isomers.

Figure 1. ISPs of the CH_3 group of both conformers ec and st (CH_3 staggering $C_4{=}C_5$ not shown in Figure 1, are collected in Table I. The remaining ISPs of the st conformers either will not

(15) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1976**, *26*, 213.
(16) Frisch, J. M. J.; Head-Gordon, M.; Trucks, G. W.; Foreman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M.; Blinkeley, S.; Gonzales, C.; Defrees, D. F.; Fox, D. J.; Whiteside, R. A.; Kahn, L. R.; Stewart, J. J. P.; Topiol. S.; Pople, J. A. *Gaussian 90*; Gaussian Inc.: Pittsburgh, PA 1990.

Table I. Thymine Isomers: ISPs of Region of CH_3 Group (Staggered Conformation)[a]

species[b]	∠6,5,C·5,HC·5 o.p.[d]	∠5,C·5,HC·5 i.p.[e]	∠5,C·5,HC·5 o.p.	d(5,C·5)	d(C·5,HC·5) i.p.	d(C·5,HC·5) o.p.	∠6,5,C·5	∠4,5,C·5	V_3[c]
Thyl ec	±119.3			1.502			117.9	124.5	1.60
Thyl st	±120.1	109.4	111.5	1.509	1.080	1.085	118.8	123.9	
Thy2 ec	±120.5			1.502			117.9	124.7	1.44
Thy2 st	±119.7	109.8	111.4	1.509	1.081	1.085	118.8	124.0	
Thy3 ec	±119.4			1.502			117.6	124.6	1.48
Thy3 st	±119.7	109.7	111.4	1.509	1.081	1.085	118.7	123.8	
Thy4a ec	±120.1			1.504			124.8	120.2	1.73
Thy4a st	±121.1	110.5	111.4	1.511	1.080	1.085	123.2	123.8	
Thy5s ec	±119.3			1.507	1.083	1.088	122.9	123.4	2.86
Thy5s st	±120.2	113.3	111.4	1.515	1.084	1.085	125.5	121.1	
Thy6 ec	±119.9	110.3	111.4	1.505	1.084	1.085	122.4	124.4	1.02
Thy6 st	±119.9	110.8	111.2	1.509	1.081	1.086	123.7	123.2	
Thy7 ec	±119.1	110.3	112.2	1.508	1.083	1.088	123.0	123.7	1.85
Thy7 st	±119.9	113.3	111.2	1.513	1.085	1.085	125.6	120.9	
Thy8 ec	±120.1	110.3	111.4	1.505	1.084	1.085	122.3	124.5	1.12
Thy8 st	±119.9	110.8	111.2	1.509	1.081	1.086	123.6	123.3	
Thy9 ec	±119.2	110.3	110.3	1.508	1.083	1.088	123.0	123.3	2.01
Thy9 st	±119.9	113.4	111.2	1.513	1.084	1.085	125.6	120.9	
Thy10a ec	±119.6	111.9	109.9	1.508	1.084	1.085	123.0	123.3	0.14
Thy10a st	±119.9	110.9	111.3	1.508	1.086	1.086	124.3	122.1	
Thy10s ec	±118.8	109.9	112.6	1.509	1.083	1.088	123.5	123.1	0.66
Thy10s st	±119.8	112.9	111.4	1.509	1.086	1.086	125.8	120.8	
Thylla ec	±121.0	111.2	110.4	1.501	1.085	1.084	117.8	123.9	1.82
Thylla st	±119.6	109.3	111.4	1.509	1.080	1.085	119.0	123.1	
Thylls ec	±121.0	111.3	110.3	1.500	1.085	1.084	118.1	123.9	1.84
Thylls st	±119.6	109.2	111.4	1.508	1.080	1.085	119.1	123.2	

[a] Angles ∠A,B,C in deg, bond length d(A,B) in angstroms. [b] Cf. Chart I. [c] V_3 = barrier to CH_3 internal rotation, kcal/mol. [d] o.p. = out-of-plane H atoms. [e] i.p. = in-plane H atom.

Table II. Thymines: Calculated Total Energies (au), Relative Energies (kcal/mol), Zero-Point Energies (kcal/mol), and Electric Dipole Moments (D)

species[a]	HF/ 6-31G** $-\mathcal{E}_t$	$\Delta\mathcal{E}$	MP2/ 6-31G** $-\mathcal{E}_t$	$\Delta\mathcal{E}$	ZPE	dipole moment
Thyl ec	451.524 19	0	452.890 03	0	77.89	3.886
Thyl st	451.521 64	1.60	452.887 93	1.31	77.71	3.924
Thyl2 ec	451.490 54	21.11	452.858 97	19.49	77.12	2.012
Thy2 st	451.488 25	22.55	452.856 86	20.96	77.00	1.974
Thy3 ec	451.505 61	11.66	452.873 08	10.64	77.78	2.442
Thy3 st	451.503 24	13.15	452.871 03	11.92	77.59	2.459
Thy4 ec	451.500 63	14.78	452.869 04	13.17	77.80	4.659
Thy4 st	451.497 87	16.52	452.866 36	14.85	77.65	4.716
Thy5 ec	451.486 13	23.88	452.855 51	21.66	77.47	7.547
Thy5 st	451.481 58	26.78	452.851 06	24.45	77.19	7.630
Thy6 ec	451.500 24	15.03	452.869 37	12.96	77.64	2.890
Thy6 st	451.498 61	16.05	452.867 86	13.91	77.45	2.791
Thy7 ec	451.491 46	20.54	452.860 89	18.28	77.50	4.084
Thy7 st	451.488 51	22.39	452.858 12	20.02	77.24	4.081
Thy8 ec	451.502 28	13.75	452.871 44	11.66	77.72	1.406
Thy8 st	451.500 49	14.87	452.869 67	12.77	77.54	1.430
Thy9 ec	451.491 37	20.59	452.861 12	18.14	77.51	4.280
Thy9 st	451.488 17	22.60	452.858 02	20.09	77.22	4.352
Thy10a ec	451.477 47	29.32	452.846 71	27.18	76.67	5.846
Thy10a st	451.477 25	29.45	452.846 45	27.34	76.74	5.766
Thy10s ec	451.482 74	26.01	452.851 74	24.02	77.14	7.092
Thy10s st	451.481 69	26.67	452.850 16	25.02	77.02	7.036
Thylla ec	451.492 23	20.05	452.859 41	19.21	77.53	5.603
Thylla st	451.489 33	21.87	452.856 83	20.83	77.37	5.661
Thylls ec	451.473 19	32.00	452.841 20	30.64	76.58	8.329
Thylls st	451.470 26	33.84	452.838 82	32.13	76.38	8.373

[a] Cf. Chart I, ec (st) symbolizes CH_3 group eclipsing (staggering) the C_4,C_5 bond.

be explicitly reported since they differ only slightly from those of the ec conformer or else they will be considered in the correlations between geometric isomerism (conformation) and ISPs. In Table II,total electronic energies of both conformers ec and st are listed together with relative energies and electric dipole moments.

Total electronic energies and energies of conversion of five cytosine and four uracil isomers (cf. Charts II and III) are given in Table III; these data are used later in the linear contribution model for conversion energies.

A selection of quantum chemical predictions of electric field gradients at ring nitrogen nuclei of thymines is listed in Table IV; it will serve below in the discussion of attractive interactions. Finally, in Table V, predicted and, where available, experimental rotational constants of the four lowest energy isomers are reported.

In the Appendix, quantum chemical energies of a small set of model compounds required for estimation of specific repulsive interactions of thymine are given.

Analysis and Discussion

Energetic Aspects of Thymine Isomers. Relative Stability. From Table II, one finds the following electronic energy (stability) sequence (cf. Chart I; ≲ symbolizes differences ≤ 0.5 kcal/mol, ec conformer)

HF/6-31G**:

Thyl < 3 < 8 < 4 ≲ 6 < 11a ≲ 7 ≲ 9 ≲ 2 < 5 < 10s < 10a < 11s

MP2/HF/6-31G**:

Thyl < 3 < 8 < 6 ≲ 4 < 9 ≲ 7 < 11a ≲ 2 < 5 < 10s < 10a < 11s

Obviously, the sequence depends on the computational level at each position, where differences in electronic energy are small. A similar assertion may be made for uracil (cf. ref 1 and Table III), where the complete sequence now reads

HF/6-31G**:

U1 < 3 < 8 < 4 ≲ 6 < 7 ≤ 9 < 11a ≲ 2 ≲ 5 < 10s < 10a < 11s

which also holds for MP2/HF/6-31G**.

Inclusion of the zero-point energy (ZPE) does not affect noticeably the stability sequences.

In all cases, the ec conformation (local minimum) is more stable than the st conformation (saddle point), but the energy difference varies widely between ≈0.1 and 2.8 kcal/mol. The

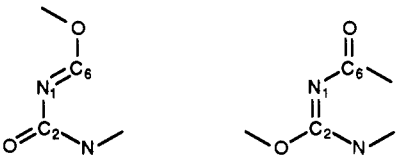
Table III. Cytosine and Uracil Isomers: Total Electronic and Geometric Isomer Conversion Energies

species[a]	HF/6-31G** $-\mathscr{E}_t$ (au)	MP2/HF/6-31G** $-\mathscr{E}_t$ (au)	ZPE[b]	HF/6-31G** $\mathscr{E}_t - \mathscr{E}_{tref}$[c]	MP2/HF/6-31G** $\mathscr{E}_t - \mathscr{E}_{tref}$[c]	$\Delta U°(0)$[d]
Cy10	392.619 59	393.827 793	66.06			
Cy11aa	392.588 85	393.799 118	66.45	0	0	
Cy11as	392.596 23	393.805 735	66.93	−4.65	−4.15	−3.67[g]
Cy11sa	392.605 66	393.814 967	67.35	−10.55	−9.95	−9.05[h]
Cy11ss	392.610 99	393.819 495	67.54	−9.26	−8.63	−8.02[i]
				−3.34	−2.84	−2.65[j]
U10a	412.438 84	413.654 372	58.35	0	0	0
U10s	412.444 79	413.659 899	58.75	−3.73	−3.46	−3.06
U11a	412.448 85	413.663 803	58.95	0	0	0
U11s	412.429 49	413.645 147[e]	58.01	12.15	11.17	10.17
	412.429 53	413.645 152[f]	58.16			

[a] Notation for isomers, cf. Schemes II and III. [b] ZPE, zero-point energy, unscaled (kcal/mol). [c] Electronic energy of geometric isomer conversion (kcal/mol). [d] Inner energy of conversion (kcal/mol). [e] Planar constraint optimization. [f] Unconstrained (nonplanar) optimized geometry. [g] Cy11aa → 11as. [h] Cy11aa → 11sa. [i] Cy11as → 11ss. [j] Cy11sa → 11ss.

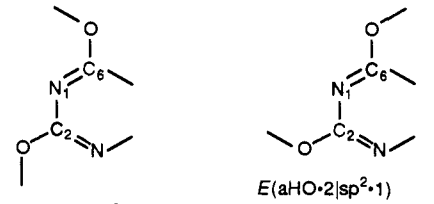
latter is closely related to the potential barrier to CH_3 rotation, which will be considered below.

Additive Contributions Model for Conversion Energies of Thymine Isomers. From the electronic energy data (Table II), the conversion energies of pairs of geometric isomers listed in Table VI are obtained. Based on this table and the conventions put forward above, one finds the equations below expressing conversion energies additively by increments (contributions). The latter should be straightforwardly perceptible from the molecular structures involved in the conversion—as the reader might perceive for himself. The equations will be formulated for HF/6-31G** data (kcal/mol), t_1, t_2, ..., t_{10} will be used as short notations for the increments. Conversion equations associated with the set of st conformers will be given separately with specific increments marked by st. Furthermore, two versions of the set of equations will be given. In version 2 (primed equation numbers), the attractive interaction associated with fragements



$E(aHO{\cdot}6|sp^2{\cdot}1||0{:}2)$ $E(aHO{\cdot}2|sp^2{\cdot}1||0{:}6)$ (F2)

will be distinguished from the ones associated with fragments (version 1)



$E(aHO{\cdot}6|sp^2{\cdot}1)$ $E(aHO{\cdot}2|sp^2{\cdot}1)$ (F1)

This distinction is not made in version 1.

Now the equations read

$$\Delta\mathscr{E}_{2\to3} \approx -9.46$$

$$\approx -E_{15}(aHO{\cdot}2|H{\cdot}1)(t_1) + E(sHO{\cdot}2|sp^2{\cdot}3)(t_2) \quad (1\text{-T})$$

$$\Delta\mathscr{E}_{2\to3} \approx -9.41$$

$$\approx -E_{15}(aHO{\cdot}2|H{\cdot}1)(t_1) + E(sHO{\cdot}2|sp^2{\cdot}3)(t_2) \quad (1\text{-T st})$$

$$\Delta\mathscr{E}_{4\to5} \approx +9.10$$

$$\approx -E(aHO{\cdot}6|sp^2{\cdot}1)(t_3) + E_{16}(sHO{\cdot}6|CH_3{\cdot}5)(t_8) \quad (2\text{-T})$$

$$\approx -E(aHO{\cdot}6|sp^2{\cdot}1|2{:}0)(t_9) + E_{16}(sHO{\cdot}6|CH_3{\cdot}5)(t_8) \quad (2\text{-T}')$$

Table IV. Electric Field Gradients at Ring Nitrogens of the Diastereomers 6–9 of Thymine[a]

	diastereomer[c]			
	Thy6 ec	Thy7 ec	Thy8 ec	Thy9 ec
	Nitrogen N_1			
q_{AA}	−0.5248	−0.5380	−0.5347	−0.5446
q_{BB}	−0.2244	−0.2880	−0.2739	−0.3369
q_{CC}	0.7492	0.8180	0.8085	0.8814
η[b]	0.401	0.315	0.323	0.236
	Nitrogen N_3			
q_{AA}	−0.5187	−0.5086	−0.5152	−0.5085
q_{BB}	−0.3750	−0.3763	−0.3078	−0.3157
q_{CC}	0.8939	0.8849	0.8230	0.8241
η[b]	0.161	0.150	0.252	0.234

[a] Atomic units, principal axis systems of field gradient tensor. [b] Asymmetry parameter. [c] Cf. Chart I for notation.

Table V. Rotational Constants (MHz) of the Four Lower Lying Tautomers of Thymine[a]

	tautomer			
	Thy 1	Thy 3	Thy 8	Thy 4
A	3257 (3256)	3283 (3293)	3230 (3227)	3278
B	1426 (1414)	1420 (1404)	1446 (1426)	1441
C	998 (992)	998 (990)	1005 (995)	1007

[a] Values in parentheses are experimental rotational constants, cf. ref 11.

$$\Delta\mathscr{E}_{4\to5} \approx +10.22$$

$$\approx -E(aHO{\cdot}6|sp^2{\cdot}1)(t_3) + E_{16}(sHO{\cdot}6|stCH_3{\cdot}5)(t_8) \quad (2\text{-T st})$$

$$\approx -E(aHO{\cdot}6|sp^2{\cdot}1|2{:}0)(t_9) + E_{16}(sHO{\cdot}6|stCH_3{\cdot}5)(t_8) \quad (2\text{-T st}')$$

$$\Delta\mathscr{E}_{6\to7} \approx +5.51$$

$$\approx -E(aHO{\cdot}2|sp^2{\cdot}1|aHO{\cdot}6)(t_4) + E(aHO{\cdot}2|sp^2{\cdot}1)(t_5) + E_{16}(sHO{\cdot}6|CH_3{\cdot}5)(t_8) \quad (3\text{-T})$$

$$\Delta\mathscr{E}_{6\to7} \approx +6.34$$

$$\approx -E(aHO{\cdot}2|sp^2{\cdot}1|aHO{\cdot}6)(t_4) + E(aHO{\cdot}2|sp^2{\cdot}1)(t_5) + E_{16}(sHO{\cdot}6|stCH_3{\cdot}5)(t_8) \quad (3\text{-T st})$$

$$\Delta\mathscr{E}_{6\to8} \approx -1.28\ (-1.18)$$

$$\approx -E(aHO{\cdot}2|sp^2{\cdot}1|aHO{\cdot}6)(t_4) + E(sHO{\cdot}2|sp^2{\cdot}3)(t_2) + E(aHO{\cdot}6|sp^2{\cdot}1)(t_3)(4\text{-T}) \quad (4\text{-T st})$$

Table VI. Thymine: Conversion Energies of Pairs of Geometric Isomers (kcal/mol)

conversion[a]	$\Delta\mathcal{E}$ HF/6-31G**	$\Delta\mathcal{E}$ MP2/HF/6-31G**	$\Delta U^{\circ}(0)$[b]
2a ec → 3s ec	−9.46	−8.85	−8.19
st → st	−9.41	−8.89	−8.30
4a ec → 5s ec	9.10	8.49	8.16
st → st	10.22	9.63	9.17
6aa ec → 7as ec	5.51	5.32	5.18
st → st	6.34	6.11	5.90
6aa ec → 8sa ec	−1.28	−1.30	−1.22
st → st	−1.18	−1.14	−1.05
7as ec → 9ss ec	0.06	−0.14	−0.13
st → st	0.21	0.06	0.04
8sa ec → 9ss ec	6.85	6.48	6.27
st → st	7.73	7.31	6.99
10a ec → 10s ec	−3.31	−3.16	−2.69
st → st	−2.79	−2.33	−2.05
11a ec → 11s ec	11.95	11.43	10.48
st → st	11.97	11.30	10.31

[a] Cf. Chart I. [b] Unscaled ZPE is used.

$$\Delta\mathcal{E}_{7\to9} \approx +0.06\ (+0.21)$$
$$\approx -E(\mathrm{aHO{\cdot}2|sp^2{\cdot}1})(t_5) + E(\mathrm{sHO{\cdot}2|sp^2{\cdot}3})(t_2)(5\text{-T})\ (5\text{-T st})$$

$$\Delta\mathcal{E}_{8\to9} \approx +6.85$$
$$\approx -E(\mathrm{aHO{\cdot}6|sp^2{\cdot}1})(t_3) + E_{16}(\mathrm{sHO{\cdot}6|CH_3{\cdot}5})(t_8)\ (6\text{-T})$$

$$\Delta\mathcal{E}_{8\to9} \approx +7.73$$
$$\approx -E(\mathrm{aHO{\cdot}6|sp^2{\cdot}1})(t_3) + E_{16}(\mathrm{sHO{\cdot}6|stCH_3{\cdot}5})(t_8)\ (6\text{-T st})$$

$$\Delta\mathcal{E}_{10a\to10s} \approx -3.31$$
$$\approx -E_{15}(\mathrm{aHO{\cdot}6|H{\cdot}1})(t_6) + E_{16}(\mathrm{sHO{\cdot}6|CH_3{\cdot}5})(t_8)\ (7\text{-T})$$

$$\Delta\mathcal{E}_{10a\to10s} \approx -2.79$$
$$\approx -E_{15}(\mathrm{aHO{\cdot}6|H{\cdot}1})(t_6) + E_{16}(\mathrm{sHO{\cdot}6|stCH_3{\cdot}5})(t_8)\ (7\text{-T st})$$

$$\Delta\mathcal{E}_{11a\to11s} \approx +11.95\ (+11.97)$$
$$\approx -E(\mathrm{aHO{\cdot}2|sp^2{\cdot}1})(t_5) + E_{15}(\mathrm{sHO{\cdot}2|H{\cdot}3})(t_7)(8\text{-T})\ (8\text{-T st})$$

$$\Delta\mathcal{E}_{11a\to11s} \approx +11.95\ (+11.97)$$
$$\approx -E(\mathrm{aHO{\cdot}2|sp^2{\cdot}1|6{:}0})(t_9) + E_{15}(\mathrm{sHO{\cdot}2|H{\cdot}3})(t_7)(8\text{-T}')\ (8\text{-T st}')$$

For version 1, the unprimed eqs 1–8 and 1 st–8st should be used for conformers with eclipsed and staggered CH_3 groups, respectively; for version 2, eqs 1, 2′, 3–7, 8′, and 1 st, 2 st′, 3 st–7 st, 8 st′ should be used. In either case, eqs 3, 4, 5, and 6 are linearly dependent (Hess's theorem). To make the eight contributions t_1–t_8 (t_1–t_9 in version 2) determinable, eqs 1–8 will be completed by the following zeroth-order estimates and approximate auxilliary equations derived from model compounds (cf. ref 1, Tables 6 and 8, kcal/mol): $E^{\mathrm{aux}}(\mathrm{sHO{\cdot}2|sp^2{\cdot}1})(t_5) \approx E^{\mathrm{aux}}(\mathrm{sHO{\cdot}2|sp^2{\cdot}3})(t_2) \approx E^{\mathrm{aux}}(\mathrm{aHO{\cdot}6|sp^2{\cdot}1})(t_3)$. For these quantities, two estimates are available, namely −4.27 and −5.62 kcal/mol; both were given equal weight in the least-squares process (eqs 9, 11, 12, and 13;

Table VII. Contributions to Geometric Isomer Conversion Energies (kcal/mol)[a]

contribution	thymine[b] (t_i)	uracil (u_i)	cytosine (c_i)
E_{15}(aHO·2|H·1)	4.91(64) ec (t_1) 5.08(70) st	4.92(59) (u_1)	4.6(5) (c_{10})
E(sHO·2|sp²·3)	−4.81(52) ec (t_2) −4.67(57) st	−4.74(48) (u_2)	−4.8(5) (c_5)
E(aHO·6|sp²·1)	−5.11(39) ec (t_3) −5.10(42) st	−4.95(35) (u_3)	
E_{15}(sHO·6|H·5)		1.66(30) (u_4)	
E(aHO·2|sp²·1|aHO·6)	−8.65(65) ec (t_4) −8.61(71) st	−8.44(59) (u_5)	
E(aHO·2|sp²·1)	−4.95(42) ec (t_5) −4.93(46) st	−4.83(39) (u_6)	−4.8(3) (c_6)
E_{15}(aHO·6|H·1)	5.16(61) ec (t_6) 5.42(67) st	5.15 (55) (u_7)	
E_{15}(sHO·2|H·3)	6.47(56) ec (t_7) 6.44(61) st	6.70(52) (u_8)	6.9(3) (c_7)
E_{16}(sHO·6|$^{ec}_{st}$CH₃·5)	2.10(37) ec (t_8) 2.97(41) st		
E(aHO·6|sp²·1|0:2)	−6.25(44) ec		
E(aHO·2|sp²·1|0:6)	−6.38(48) st (t_9)	−6.12(40) (u_9)	
E_{15}(aHN:6|H·1)			0.6–1.2(6) (c_1)
E_{15}(sHN:6|H·5)			2.0(2) (c_2)
E(aHO·2|sp²·1|aHN:6)			−7.3(3) (c_3,c'_3)
E(aHN:6|sp²·1)			−3.0(2) (c_4,c'_4)
E(H·1|sp²N:6)			−3.0(5) (c_8)
E_{1515}(aHO·2|H·1|aHN:6)			7.4(10) (c_9)
s_{y-x}[c]	0.62 0.68–0.71	0.57	0.4–0.7
$\bar{v}v$	3.05–4.04	3.0	4–6

[a] Values in parenthesis = estimate of rms deviation. [b] ec (st) denotes eclipsed (staggered) CH_3 conformation; for st, version 2, $b_{16} \approx +3.06$ kcal/mol has been chosen, cf. text. [c] rms deviation of residual sum of squares. [d] Total sum of squares of errors.

$$E_{15}^{\mathrm{aux}}(\mathrm{aHO{\cdot}2|H{\cdot}1})(t_1) \approx E_{15}^{\mathrm{aux}}(\mathrm{aHO{\cdot}6|H{\cdot}1})(t_6) \quad (10)$$

$$-E^{\mathrm{aux}}(\mathrm{aHO{\cdot}2|sp^2{\cdot}1|aHO{\cdot}6})(t_4) + E^{\mathrm{aux}}(\mathrm{aHO{\cdot}2|sp^2{\cdot}1})(t_5) \approx +4.00 \quad (14)$$

$$-E^{\mathrm{aux}}(\mathrm{aHO{\cdot}2|sp^2{\cdot}1})(t_5) + E_{15}^{\mathrm{aux}}(\mathrm{sHO{\cdot}2|H{\cdot}3})(t_7) \approx +12.20 \quad (15)$$

$$E_{16}^{\mathrm{aux}}(\mathrm{sHO{\cdot}6|ecCH_3{\cdot}5})(t_8) \approx +2.24 \quad \text{(only one estimate available)} \quad (16)$$

$$E_{16}^{\mathrm{aux}}(\mathrm{sHO{\cdot}6|stCH_3{\cdot}5})(t_8) \approx 3.06 \quad \text{(2.24 from model compounds, see Appendix)} \quad (16\ \mathrm{st})$$

$$E^{\mathrm{aux}}(\mathrm{aHO{\cdot}6|sp^2{\cdot}1|2{:}0})(t_9) \approx E^{\mathrm{aux}}(\mathrm{aHO{\cdot}2|sp^2{\cdot}1|6{:}0})(t_9) \approx -6.27 \quad (17)$$

From this body of information there results a linear system $At = b$ with eight (nine in version 2) unknowns and 17 equations. The solutions are collected in Table VII, including root mean square (rms) estimates of the increments. Since version 2 reproduces the conversion energies with a significantly smaller residual sum of squares, only these estimates of contributions are given. Version 1 produces values of corresponding increments very similar to those reported earlier for uracil.[1] By version 2, not only the sum of residual errors but also, as a rule, rms estimates of almost all contributions and residual errors associated with eqs 1-T–8-T are significantly reduced. This may be taken as a justification for introduction of a specific (attractive) interaction associated with fragment F2, $E(\mathrm{aHO{\cdot}2|sp^2{\cdot}1\|0{:}6}) \approx E(\mathrm{aHO{\cdot}6|sp^2{\cdot}1\|0{:}2}) \approx -6.3(5)$ kcal/mol, significantly higher in magnitude than the increments associated with structural fragment F1, e.g., $E(\mathrm{aHO{\cdot}6|sp{\cdot}1}) \approx -4.8(4)$ kcal/mol.

Some further comments concerning the contribution terms shown in Table VII seem to be appropriate.

(i) The various repulsive interactions of type E_{15} are widely different. Whereas $E_{15}(\mathrm{sHO{\cdot}6|H{\cdot}5}) \approx 1.7(3)$ kcal/mol is ap-

proximately the same as the values found for acyclic protonated carbonyl compounds,[2] the repulsive interactions of type E_{15}(aHO·2|H·1), E_{15}(aHO·6|H·1), and E_{15}(sHO·2|H·3) are much higher (5.0(6) and 6.5(6) kcal/mol, respectively).

(ii) The "double" attractive interaction ("double" H-bond) E(aHO·2|sp²·1|aHO·6) (–8.7(7) kcal/mol) is but slightly lower in magnitude than the sum of two single attractions E-(aHO·2|sp²·1) and E(aHO·6|sp²·1).

(iii) Both methyl group conformations ec and st show the same contribution terms, except eventually E_{16}(sHO·6|CH_3·5), which seems tendentially higher for st. Steric crowding might be a straightforward interpretation for this observation (see also the section on structural relaxation).

Updating of Additive Increment Models for Uracil and Cytosine. For uracil, determination of contributions to conversion energies was earlier based on isomers U2–U9. Use of the new computations for U10a, U10s, U11a, and U11s (cf. Chart 3) affords both improvement and extension of the earlier model and, in addition, a nearly independent check of the data for thymine. In fact, corresponding increments will be found equal within error limits.

The isomers U2–U11s provide for eqs 1-U–8-U, which will again be formulated for versions 1 and 2 (unprimed and primed equations in place of unprimed equations, respectively). In version 2, which is analogous to version 2 of the thymine case, the attractive interactions associated with fragments F1 and F2 will be distinguished. Using HF/6-31** data, the equations for the linear increments read (u_1–u_8 and u_1–u_9 serving as short notations, energies in kcal/mol)

$$\Delta\mathscr{E}_{2\to3} \approx -9.42$$
$$\approx -E_{15}(\text{aHO·2|H·1})(u_1) + E(\text{sHO·2|sp}^2\text{·3})(u_2) \quad (1\text{-U})$$

$$\Delta\mathscr{E}_{4\to5} \approx +8.32$$
$$\approx -E(\text{aHO·6|sp}^2\text{·1})(u_3) + E_{15}(\text{sHO·6|H·5})(u_4) \quad (2\text{-U})$$
$$\approx -E(\text{aHO·6|sp}^2\text{·1‖0:2})(u_9) + E_{15}(\text{sHO·6|H·5})(u_4) \quad (2\text{-U}')$$

$$\Delta\mathscr{E}_{6\to7} \approx +4.82$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1|aHO·6})(u_5) + E(\text{aHO·2|sp}^2\text{·1})(u_6) + E_{15}(\text{sHO·6|H·5})(u_4) \quad (3\text{-U})$$

$$\Delta\mathscr{E}_{6\to8} \approx -1.20$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1|aHO·6})(u_5) + E(\text{sHO·2|sp}^2\text{·3})(u_2) + E(\text{aHO·6|sp}^2\text{·1})(u_2) \quad (4\text{-U})$$

$$\Delta\mathscr{E}_{7\to9} \approx +0.10$$
$$\approx =E(\text{aHO·2|sp}^2\text{·1})(u_6) + E(\text{sHO·2|sp}^2\text{·3})(u_2) \quad (5\text{-U})$$

$$\Delta\mathscr{E}_{8\to9} \approx +6.12$$
$$\approx -E(\text{aHO·6|sp}^2\text{·1})(u_3) + E_{15}(\text{sHO·6|H·5})(u_4) \quad (6\text{-U})$$

$$\Delta\mathscr{E}_{10a\to10s} \approx -3.73$$
$$\approx -E_{15}(\text{aHO·6|H·1})(u_7) + E_{15}(\text{sHO·6|H·5})(u_4) \quad (7\text{-U})$$

$$\Delta\mathscr{E}_{11a\to11s} \approx +12.15$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1})(u_6) + E_{15}(\text{sHO·2|H·3})(u_8) \quad (8\text{-U})$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1‖0:6})(u_9) + E_{15}(\text{sHO·2|H·3})(u_8) \quad (8\text{-U}')$$

By Hess's theorem, eqs 3-U–6-U are linearly dependent. Determination of increments u_1–u_8 (u_1–u_9 for version 2) is enforced by using the following zeroth-order estimates derived from model compounds (cf. ref 1 and thymine above, kcal/mol): E_{15}(aHO·2|H·1)(u_1) ≈ E_{15}(aHO·6|H·1)(u_7) and E(sHO·2|sp²·3)(u_2) ≈ E(aHO·2|sp²·1)(u_6) ≈ E(aHO·6|sp²·1)(u_3) (for these quantities, two estimates, –4.27 and –5.62 kcal/mol, are available from model compounds, each of which will be used with weight 1); E^{aux}_{15}(sHO·6|H·5)(u_4) ≈ +1.98 (this estimate originates from model compounds for acyclic imines and O-protonated carbonyl compounds and is given weight 2 (cf. refs 2 and 3); and $-E^{aux}$(aHO·2|sp²·1|aHO·6)(u_5) + E^{aux}(aHO·2|sp²·1)(u_6) ≈ +4.00 and $-E^{aux}$(aHO·2|sp²·1)(u_6) + E^{aux}_{15}(sHO·2|H·3)(u_8) ≈ +12.20. For version 2, eq 2-U′ in place of 2-U and the auxilliary condition E^{aux}(aHO·6|sp²·1‖0:2) ≈ E^{aux}(aHO·2|sp²·1‖0:6) ≈ –6.27 will be used.

In total there are then available in versions 1 and 2 16 conditions for unknowns u_1–u_8 and 17 conditions for unknowns u_1–u_9, respectively, forming linear systems of type $Au = b$. Solution by the least-squares method shows version 2 to give nearly half the residual sum of squares and significantly reduced errors of eqs 1-U–8-U and reduced estimated rms deviations of the increments as compared to version 1. In both cases, the normal matrix is well conditioned. The least-squares estimates of the former are collected in Table VII. Comparison with the values for thymine shows corresponding contributions to agree within error estimates. Version 1 yields contributions esentially equal to the values reported earlier (ref 1, Table 8). The improvements achieved by version 2 mostly stem from introduction of a specific increment ascribed to substructure F2.

In the case of cytosine, the newly studied isomers Cyllaa–Cyllss supply four (three linearly independent) new equations to the set of six (five linearly independent) equations associated with isomers Cy2–Cy9 studied earlier. For the sake of convenience, the full set of equations will be reproduced (cf. Table 5 of ref. 1, and Table III above). Data from HF/6-31G** computations (kcal/mol) are used, and c_1, ..., c_{10} serve as short notations for increments:

$$\Delta\mathscr{E}_{2\to3} \approx -1.71$$
$$\approx -E_{15}(\text{aHN:6|H·1})(c_1) + E_{15}(\text{sHN:6|H·5})(c_2) + E(\text{H·1|sp}^2\text{N:6})(c_8) \quad (1\text{-Cy})$$

$$\Delta\mathscr{E}_{4\to5} \approx -0.75$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1|aHN·6})(c_3) + E(\text{aHN·6|sp}^2\text{·1})(c_4) + E(\text{sHO·2|sp}^2\text{·3})(c_5) \quad (2\text{-Cy})$$

$$\Delta\mathscr{E}_{6\to7} \approx +4.34$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1|aHN:6})(c'_3) + E(\text{aHO·2|sp}^2\text{·1})(c_6) + E_{15}(\text{sHN:6|H·5})(c_2) \quad (3\text{-Cy})$$

$$\Delta\mathscr{E}_{6\to8} \approx +11.15$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1|aHN:6})(c'_3) + E(\text{aHN:6|sp}^2\text{·1})(c'_4) + E_{15}(\text{sHO·2|H·3})(c_7) \quad (4\text{-Cy})$$

$$\Delta\mathscr{E}_{7\to9} \approx +12.20$$
$$\approx -E(\text{aHO·2|sp}^2\text{·1})(c_6) + E_{15}(\text{sHO·2|H·3})(c_7) \quad (5\text{-Cy})$$

$$\Delta\mathscr{E}_{8\to9} \approx +5.39$$
$$\approx -E(\text{aHN:6|sp}^2\text{·1})(c'_4) + E_{15}(\text{sHN:6|H·5})(c_2) \quad (6\text{-Cy})$$

$$\Delta\mathscr{E}_{llaa\to llas} \approx -4.65$$
$$\approx -E_{1515}(\text{aHO·2|H·1|aHN:6})(c_9) + E_{15}(\text{aHO·2|H·1})(c_{10}) + E_{15}(\text{sHN:6|H·5})(c_2) + E(\text{H·1|sp}^2\text{N:6})(c_8) \quad (7\text{-Cy})$$

Table VIII. ISP–Geometric Isomer Correlations of Thymine Isomers Thy–9

ISP	regression coefficients[a] a_2	a_6	a_5	ISP	regression coefficients[a,b] a_2	a_6	a_5
∠1,2,3	0.09(2)	0.39(2)	−0.06(2)	∠1,2,O·2	−1.1(1)	0	0.2(0)
∠2,3,4	0.5(1)	−0.5(1)	−0.11(8)	∠3,2,O·2	1.0	−0.5	−0.2
∠3,4,5	−0.7(1)	0.2(1)	0.1(1)	∠2,O·2,HO·2	−0.5(1)	−0.2(1)	0
∠5,6,1	0.5(1)	−0.7(1)	−0.3(1)	*d*(2,O·2)	0	−1	0
<6,1,2	−0.4(1)	0.3(1)	0.4(1)	*d*(O·2,HO·2)	0	0.8(2)	0
d(1,2)	−4(1)	−5(1)	−4(1)	∠1,6,O·6	0	−2.8(1)	−0.7(1)
d(2,3)	4(1)	5(1)	3(1)	∠5,6,O·6	−1.0(0)	3.5(1)	1.0(1)
d(3,4)	4(1)	−7(1)	−4(1)	∠6,O·6,HO·6	0	3.2	0
d(4,5)	−4(1)	4(1)	4(1)	*d*(6,O·6)	−1.0(2)	2.8	0.8(2)
d(5,6)	5(1)	4(1)	−2(1)	*d*(O·6,HO·6)	0	−4.5	−0.5
d(6,1)	−3(1)	−1(1)	−4(1)	∠4,5,C·5	0	−0.9(1)	−1.9(1)
				∠6,5,C·5	0	1.1(3)	1.8(3)
				d(5,C·5)	0	3.4(2)	4.4

[a] Units mÅ and deg for bond length and bond angles, respectively. [b] Cf. eq 1.

$$\Delta\mathcal{E}_{\text{llaa}\rightarrow\text{llsa}} \approx -10.56$$

$$\approx -E_{1515}(\text{aHO·2}|\text{H·1}|\text{aHN:6})(c_9) + E_{15}(\text{aHN:6}|\text{H·1})(c_1) + E(\text{sHO·2}|\text{sp}^2\text{·3})(c_5) \quad (8\text{-Cy})$$

$$\Delta\mathcal{E}_{\text{llas}\rightarrow\text{llss}} \approx -9.25$$

$$\approx -E_{15}(\text{aHO·2}|\text{H·1})(c_{10}) + E(\text{sHO·2}|\text{sp}^2\text{·3})(c_5) \quad (9\text{-Cy})$$

$$\Delta\mathcal{E}_{\text{llsa}\rightarrow\text{llss}} \approx -3.34$$

$$\approx -E_{15}(\text{aHN:6}|\text{H·1})(c_1) + E(\text{H·1}|\text{sp}^2\text{N:6})(c_8) + E_{15}(\text{sHN:6}|\text{H·5})(c_2). \quad (10\text{-Cy})$$

Among these equations, those of the quadruplets 6–9 and 7–10 are linearly dependent. To arrive at a solution for the unknown contributions c_1–c_{10}, c'_3, and c'_4, the following auxilliary assumptions and data, applied in part already earlier, are used (HF/6-31G** values, kcal/mol):

$$c_3 = c'_3 \text{ (eqs 2-Cy–4-Cy)} \quad \text{and}$$

$$c_4 = c'_4 \text{ (eqs 2-Cy, 4-Cy, and 6-Cy)} \quad (1)$$

(this reduces the number of unknown increments to 10);

$$-E^{\text{aux}}_{-15}(\text{aHN:6}|\text{H·1})(c_1) + E^{\text{aux}}(\text{H·1}|\text{sp}^2\text{N:6})(c_8) = -3.32 \quad (2)$$

(a second estimate for the right-hand side is −2.48, which, however, should be given significantly lower weight):

$$E^{\text{aux}}_{15}(\text{aHN:6}|\text{H·5})(c_2) \approx +1.99 \quad (3)$$

$$-E^{\text{aux}}(\text{aHO·2}|\text{sp}^2\text{·1}|\text{aHN:6})(c_3) + E^{\text{aux}}(\text{aHO·2}|\text{sp}^2\text{·1})(c_6) \approx +2.37 \quad (4)$$

$$E_a{}^{\text{aux}}(\text{aHN·6}|\text{sp}^2\text{·1})(c_4) \approx -2.75 \quad (5)$$

$$E^{\text{aux}}(\text{sHO·2}|\text{sp}^2\text{·3})(c_5) \approx E^{\text{aux}}(\text{aHO·2}|\text{sp}^2\text{·1})(c_6) \quad (6)$$

(for these two quantities, two estimates are available; in the case of cytosine the second should be given higher weight (see above for uracil));

$$-E^{\text{aux}}(\text{aHO·2}|\text{sp}^2\text{·1}|\text{aHN:6})(c_3) + E^{\text{aux}}(\text{aHN·6}|\text{sp}^2\text{·1})(c_4) \approx +3.53 \quad (7)$$

$$E^{\text{aux}}_{15}(\text{sHO·2}|\text{H·3})(c_7) \approx +6.6 \quad (8)$$

For $E^{\text{aux}}(\text{aHO·2}|\text{sp}^2\text{·1}|\text{aHO:6})(c_3)$, a range of −6.6 to −8.0 is available;

however, no specific choice can be proposed. (9)

For $E^{\text{aux}}_{15}(\text{aHO·1}|\text{H·1})$, an estimate of +4.68 from model compounds or

+4.92 from uracil may be chosen. (10)

Besides estimate 8 borrowed from uracil data, all estimates are provided for by model compounds. Maximum use of these leads to a linear system of 22 equations for the unknowns c_1–c_{10}. Least-squares solutions were determined by a wide variety of open choices allowed by eqs 2, 6, and 9. In Table VII, the most plausible solution is reproduced. Some complementary comments regarding the energy contributions should be made.

(i) Cytosine increments generally have higher error estimates; this in part goes back to the larger number of increments required for this system and to a less well conditioned least-squares problem. The latter as a whole is less well balanced wrt the set of unknowns than in the case of uracil and thymine.

(ii) Contributions occurring in common with uracil and thymine are equal within error estimates.

(iii) The attractive interactions E(aHN:6|sp²·1) and E(H·1|sHN:6), in which donor and acceptor of the "intramolecular" H-bridge are interchanged, are essentially equal but are absolutely lower than attractions of type E(aHO·2|sp²·1) and E(sHO·2|sp²3) (−3.0 vs −4.8 kcal/mol).

(iv) Whereas the "doubly attractive" contribution E(aHO·2|sp²·1|aHN:6) ≈ −7.3 (+4, −0) kcal/mol essentially equals the sum E(aHO·2|sp²·1) + E(aHN:6|sp²·1) ≈ −7.85 kcal/mol, the "double repulsive" interaction E_{1515}(aHO·2|H·1|aHN:6) ≈ +7.4(10) kcal/mol deviates significantly from the sum of E_{15}(aHO·2|H·1) and E_{15}(aHN:6|H·1) (+4.6(5) and 0.6–1.2(6) kcal/mol, respectively). However, both E_{1515} and E_{15}(aHN:6|H·1) are not well determined.

(v) By the additional data from Cyllaa–Cyllss, the accuracy of the increments is not significantly improved, in contrast to the situation with uracil and thymine; the reason lies mostly in more complex condition equations for Cyll conversions and unfavorable conditions for the least-squares model. For practical purposes, the increments are believed to be satisfactorily reliable.

Correlation between ISPs and Geometric Isomerism. Earlier work has shown that *alterations* of predicted ISPs brought about by interconversion of geometric isomers are nearly independent of the quantum chemical approximation (HF/6-31G* or higher).[2,3] The following consideration, therefore, is restricted to *changes* of ISPs in the thymine set. Analysis of all structural data given in Figure 1 according to the linear regression formula 1 yields similar regression coefficients as observed with cytosine and uracil for corresponding ISPs. Mainly, correlations including noticeable effects of methyl group conformational changes will be discussed.

First consider correlations (structural relaxation) of anti–syn conversions of 2- and 6-OH groups with ISPs of the regions of the two substituents within the quartet of geometric isomers Thy6–9. Part of the results is collected in Table VIII; in the left side,

Table IX. Structure–Geometric Isomerism Correlations for All Thymine Isomers except Thy6–9

		regression coefficient[a]				regression coefficient[a]		
isomer	ISP	a_2	a_6	a_5	isomer	a_2	a_6	a_5
Thy2–3	∠1,2,O·2	–3.7(1)		0	Thy11as	–2.6		0
	∠3,2,O·2	2.7(1)		0		3.5		0
	∠2,O·2,HO·2	–5.1(1)		0		5.9(1)		0
	d(2,O·2)	–2.7(9)		0		6.0(0)		0
	d(O·2,HO·2)	4.7(5)		0		–5.0(1)		0
	∠4,5,C·5	0		–0.7(1)		0		–0.8(1)
	∠6,5,C·5	–0.2(1)		1.0		0.2(1)		1.1(1)
	d(5,C·5)	0		7.0(1)		–1.0(1)		8.0
Thy4–5	∠1,6,O·6		–2.5(1)	–0.5(1)	Thy10as		–5.3(1)	–0.7(1)
	∠5,6,O·6		3.3(1)	0.8(1)			5.0(2)	0.9(2)
	∠6,O·6,HO·6		4.4(2)	0			–1.9	0
	d(6,O·6)		4.7(5)	0			–1.0	2.0
	d(O·6,HO·6)		–7.0(1)	–1.0			0	0
	∠4,5,C·5		–1.8(6)	–1.4(6)			–0.6(5)	–1.6(5)
	∠6,5,C·5		1.4(1)	1.7(1)			0.8(1)	1.6(1)
	d(5,C·5)		3.3(4)	7.3(4)			1.0(1)	0

[a] Regression coefficient in mÅ and deg, cf. eq 1.

regression coefficients of some ISPs of the ring, and on the right side, regression coefficients in the region of the OH substituents are collected. These data should be completed by the following comments.

(i) Ring bond angles vary but slightly (≤0.7°) upon anti–syn or ec–st conversion. In cases where the regression coefficients a_2 or a_6 are significantly nonzero, they have different sign, i.e., ring angles are anticorrelated wrt anti–syn conversion of 2-OH and 6-OH geometric isomers. Most ring angles relax only slightly upon ec–st conversion of the CH_3 conformation. As an exception, angle ∠6, 1, 2 should be noted.

(ii) Ring bond lengths relax surprisingly strongly upon anti–syn and ec–st conversions, i.e., by amounts of the order of ≲5 mÅ. Correlation and anticorrelation occurs wrt both types of conversion, and the coefficients are of similar magnitude. Hence, ring bond lengths afford examples of nonlocal structural relaxation. However, it seems not possible to interpret the regression coefficients in terms of repulsive and attractive interactions, in contrast to local ISPs. Furthermore, the contrast to structural relaxation of acyclic O-protonated carbonyl compounds should be noted, where vertex angles at the substituent sites are essentially equal for anti and syn isomers.

(iii) The local bond angle pairs ∠1,6,O·6,∠5,6,O·6 and ∠4,5,C·5,∠6,5,C·5 show the expected behavior: both are anticorrelated pairs wrt and anti–syn and ec–st conversion. The pair ∠1,2,0·2,∠3,2,O·2 is anticorrelated but exhibits small relaxation only.

(iv) Structural relaxation of the ISPs of the OH group in the 2-position (d(2,O·2), d(O·2,HO·2) and ∠2,O·2,HO·2) is small wrt anti–syn conversion of both the 2- and 6-OH groups and ec–st conversion of 5-CH_3. For d(6,O·6), d(0·6,HO·6), and ∠6,O·6,HO·6, dependence on geometric isomerism of the 2-OH group is rather small; however, all three ISPs depend sensitively on 6-OH isomerism: on anti–syn conversion, the two bond lengths alter by +2.8 and –4.5 mÅ, respectively, and ∠6,O·6,HO·6 increases by 3.2° (a_6 coefficients, see Table VIII). Steric crowding in the syn isomer may be responsible for these alterations: the conversion implies the repulsive interaction E_{16}(sHO·6|CH_3·5) to be switched on and the attraction E(aHO·6|sp^2·1) to be switched off. Both processes mechanically tend to enlarge ∠6,O·6,HO·6 cooperatively.

(v) For the distance d(5,C·5), a surprisingly large increase upon anti–syn and ec–st conversion of 6-OH (a_6 = 3.4 mÅ) and 5-CH_3 (a_5 = 4.4 mÅ) is observed, pictorially caused again by switching on the repulsions E_{16}(sHO·2|ecCH_3·5) and E_{16}(sHO·2|stCH_3·5). Widening and shrinking of ∠6,5,C·5 and ∠4,5,C·5, respectively, fit into this mechanical picture.

Similar correlations are found with the isomer pairs (Thy2,-Thy3), (Thy4,Thy5), (Thy10a,Thy10s), and (Thy11a,Thy11s) and their CH_3 conformers ec and st. Table IX gives a collection of regression coefficients in the more prominent cases. One should note the analogous (antianalogous) positions of hydroxy and oxo groups (and of the H substituent on N in 1- and 3-positions) in the pairs (Thy2,3)(Thy11a,11s) and (Thy4,5)(Thy10a,10s).

A few comments illustrating the interplay of attractive/repulsive interactions with alterations of local ISPs appear to be in order.

(i) The simplest behavior is shown by the local parameters of the 2-OH group, ∠2,O·2,HO·2, d(2,O·2), and d(O·2,HO·2) of the former two pairs. Anti–syn conversion Thy2→3 implies removal of repulsion E_{15}(aHO·2|H·1) and activation of the attraction E(sHO·2|sp^2·3). Both processes tend to decrease ∠2,O·2,HO·2, in agreement with the value of $a_2 \approx -5.1°$. In contrast, in the anti–syn conversion Thy11a→11s, attraction E(aHO·2|sp^2·1) is switched off and repulsion E_{15}(sHO·2|H·3) is switched on. Both changes tend to increase ∠2,O·2,HO·2, qualitatively agreeing with the regression coefficient $a_2 \approx +5.9°$. Analogous arguments predict increase and decrease of d-(O·2|HO·2) for anti–syn conversion Thy2→3 and Thy11a→11s, qualitatively interpreting the calculated a_2 values +4.7 and –5.0 mÅ, respectively. Likewise for d(2,O·2), qualitatively the opposite behavior is predicted; this should be compared to calculated a_2 values –2.7 and +6.0 mÅ. In all the foregoing examples, disappearing and newly arising interactions affect each particular ISP in the same sense.

(ii) Examples of a more delicate interplay of interactions are afforded by the local ISPs ∠1,2,O·2,∠3,2,O·2 of Thy2–3,Thy11a-11s and ∠1,6,O·6,∠5,6,O·6 of Thy4-5,Thy10a-10s, respectively. Thy2→3 conversion implies cancellation of the repulsion E_{15}-(aHO·2|H·1) (≈+5.0 kcal/mol) and creation of the attraction E(sHO·2|sp^2·3) (≈–4.7 kcal/mol), tending to shrink (increase) and to open (decrease) simultaneously ∠1,2,O·2 (∠3,2,O·2). Since the calculated regression coefficient is $a_2 \approx -3.7°$ ($a_2 \approx +2.7°$), one has to conclude that the repulsive interaction more effectively affects these ISPs. An analogous conclusion holds for HO·2 anti–syn isomerization of Thy11, though in this process the E(aHO·2|sp^2·1) attraction is removed and the E_{15}(sHO·2|H·3) repulsion is created.

For the Thy4–5 and Thy10a–10s pairs, the ISPs ∠1,6,O·6, <5,6,O·6 experience similar changes upon anti–syn conversion of the 6-OH group, as expressed by a_6 = –2.5,+3.3 and –5.3,+5.0°, respectively. For Thy4–5, the attraction E(aHO·6|sp^2·1) is set off and the repulsion E_{16}(sHO·6|CH_3·5) is set on, causing an increase of ∠1,6,O·6 (∠5,6,O·6) and decrease of ∠1,6,O·6 (∠5,6,O·6), respectively. Obviously the repulsive interaction is dominating the total change. For Thy10a–10s, anti–syn conversion removes E_{15}(aHO·6|H·1) and installs E_{16}(sHO·6|CH_3·5),

Table X. Thymine (Thy1): Structural Data Base[a]

ISP[b]	X67[d]	X61[e]	HF/6-31G** [c] st	HF/6-31G** [c] ec
d(1,2)	1.344	1.361	1.369	1.370
d(2,3)	1.313	1.354	1.367	1.368
d(3,4)	1.409	1.381	1.376	1.377
d(4,5)	1.369	1.348	1.331	1.330
d(5,6)	1.475	1.447	1.473	1.471
d(6,1)	1.412	1.390	1.387	1.387
d(2,O:2)	1.245	1.233	1.195	1.195
d(5,C·5)	1.520	1.502	1.509	1.502
d(6,O:6)	1.192	1.231	1.195	1.195
∠1,2,3	117.8	115.2	113.3	113.2
∠2,3,4	123.2	122.9	123.3	123.4
∠3,4,5	120.2	121.8	123.3	123.0
∠4,5,6	118.6	118.2	117.3	117.6
∠5,6,1	114.3	115.6	114.8	114.8
∠6,1,2	125.9	126.3	128.0	128.0
∠1,2,O:2	120.5	122.1	123.5	123.6
∠3,2,O:2	121.6	122.7	123.2	123.2
∠4,5,C·5	122.4	122.8	123.9	124.5
∠6,5,C·5	119.0	119.0	118.8	117.9
∠5,6,O:6	124.4	126.1	125.1	124.6
∠1,6,O:6	121.3	118.3	120.1	120.6

[a] Only data involving C, N, and O atoms considered; X-ray data indicate essentially planar structures. [b] Internal structural parameters (ISPs): distances d(i,k) in Å, bond angles ∠i,k,C in deg; for numbering, see Chart I. [c] This work. [d] Reference 17. [e] Reference 18.

both causing ∠1,6,O·6 to decrease (∠5,6,O·6 to increase), in qualitative agreement with $a_6 \approx -5.3°$ ($a_6 \approx +5.0°$).

(iii) The regression coefficients a_2 of d(2,O·2) of Thy2→3 and Thy11a→s (−2.7 and 6 mÅ, respectively) and those of d(O·2,-HO·2) (+4.7 and −5 mÅ, respectively) follow qualitatively the same arguments. For d(O·2,HO·2), the involved interactions are counteroperative, and the attractive term is dominating a_2; for d(2,O·2) they are cooperative. A similar interpretation may be given for the regression coefficients a_6 of d(6,O·6) and d(O·6,-HO·6) in the anti–syn conversion of the 6-OH group of Thy4–5 and Thy10a–10s, cf. Table IX.

(iv) For Thy 2–3 and Thy 11a–11s, the local ISPs ∠4,5,-C·5,∠6,5,C·5 of the CH_3 substituent in the 5-position are nearly independent of the geometric isomerism of HO·2 but sensitively alter on changing the CH_3 group conformation (cf. Table IX); this in particular applies to d(5,C·5) (a_5 = 7.0 and 8.8 mÅ, respectively).

For Thy4–5, d(5,C·5) varies noticeably on anti–syn conversion of HO·6 and ec–st change of CH_3 conformation, for Thy 10a–s, alterations are noticeably smaller in magnitude, and $a_5 \approx 0$. The latter finding is rather surprising. In these cases, ∠4,5,C·5 and ∠6,5,C·5 relax only slightly and in a rather intransparent manner. It should be kept in mind that Thy10 features a formal π electron system which is unique in the whole set of thymine isomers.

Similar remarks apply to correlations of "inner" ISPs of the methyl group proper with geometric isomerism of OH substituents and CH_3 conformation. As revealed by Table I, both local and inner ISPs vary considerably over the whole set of isomers. However, their relaxation upon anti–syn and ec–st seemingly does not follow a scheme which could be explained in terms of a typical interaction.

Comparison with Empirical Structural Data

Thymine Structural Data Base. Empirical structural data of thymine are available from two X-ray studies;[17,18] one of these has been carried out with thymine monohydrate. Both data sets are contrasted in Table X with HF/6-31G** quantum chemical predictions. Obviously, all three sets deviate significantly from each other; this is born out by comparing the means of absolute differences of corresponding ISPs of all pairs of data sets (bond length and bond angles treated separately) by the aid of the t test.[19] In all cases, empirical t values show that the data sets are different at levels of confidence >0.99. The same holds for the sets corresponding to ISPs of uracil and thymine.

Taking all presently available structural data, it should be stated that discrepancies between empirical ISPs of different provenience and quantum chemically predicted ISPs for all three pyrimidine derivatives are significant, preventing any meaningful investigation into the origin of systematic deviations of predicted ISPs for these species. The two C:O bonds are predicted to have essentially the same length (≈1.195 Å), whereas the X-ray studies yielded either different (1.245, 1.192 Å) or essentially equal (1.233, 1.233 Å) lengths, to mention just one example.

Rotational Constrants. In this work, rotational constants of all thymine isomers have been calculated, but only those of the lowest energy isomers Thy1, Thy3, Thy8, and Thy4 have been listed in Table V. For the first three species, rotational constants have been determined by microwave spectroscopy,[11] and these are likewise included in the table. Comparison shows the HF6-31G** values for A and C constants to agree within ≤10 MHz and for B constants within ≤20 MHz. The discrepancies may be considered small enough to allow useful prediction of microwave transitions with sufficient reliability, in particular if electric quadrupole hyperfine pattern based on calculated electric field gradients are included. A systematic study of this aspect of the pyrimidine nucleic acid bases is under work.

Electric Dipole Moments. Experimental dipole moments have been reported only for isomer Thy1, ranging between 3.95 and 4.13 D.[20,21]

In this work, a value of 3.886 D is predicted (cf. Table II). Previous theoretical values range between 4.86 and 5.65 D, considerably higher than experimental and present theoretical values. Over the whole set of thymine isomers, the dipole moments vary widely from 1.4 to 8.4 D; as expected, the conformation of the CH_3 group exerts nearly no influence. Pairs of geometric isomers either may possess similar moments, e.g., Thy2, Thy3 (1.99 vs 2.45 D) or Thy8, Thy9 (4.08 vs 4.32 D) or may exhibit noticeably different dipoles, e.g., Thy4, Thy5 (4.68 vs. 7.58 D) or Thy11a, Thy11s (5.63 vs 8.35 D). Factor analysis[22] indicates that simple vector addition models do not yield a satisfactory interpretation of the quantum chemical data for thymine isomers. However, this problem will be reconsidered later for the full set of pyrimidine nucleic acid bases.

Characterization of Attractive Contributions. Attractive contributions (interaction terms of type E(aHO·2|sp²·1) etc.) were found indispensable above and earlier[1] for rational interpretation of reaction energy and structural relaxation associated with anti–cis conversions of geometric isomers of some pyrimidine nucleic acid bases. Such interactions are similar to those of intramolecular hydrogen bonds in certain aspects (e.g., energy) but deviate markedly in others (e.g., geometry). Besides geometry, two further aspects will be briefly considered, namely, electronic charge density and field gradients at the nitrogen nuclei of the thymine ring.

In the earlier work, detailed information on the *structural fragments* associated with attractive interactions was given (cf. ref 1, Table 13). With uracil, the donor atom O, the H atom, and the acceptor atom N form a nearly isosceles triangle with d(N,O) ≈ 0.947(2) Å, d(O,N) ≈ 2.25(1) Å, and d(N,H) ≈ 2.23–2.29 Å. Nearly identical ISPs were obtained for thymine.

(17) Ozeki, K.; Sakabe, N.; Tanaka, J. *Acta Crystallogr., Sect. B* **1969**, *25*, 1038.
(18) Gerdil, R. *Acta Crystallogr.* **1961**, *14*, 333.
(19) Crow, E. L.; Davis, F. A.; Maxfield, M. W. *Statistics Manual*; Dover Publications: New York, 1960.
(20) Kulakowska, I.; Geller, M.; Lesyng, B.; Wierzchowski, K. L. *Biochim. Biophys. Acta* **1974**, *361*, 119.
(21) Mauret, P.; Fayet, J.-P. C. R. Hebd. *Seances Acad. Sci.* **1967**, *264*, 2081. Cited in McClellan, A. L. *Tables of Experimental Dipole Moments*; Rahara Enterprises: El Cerrito, CA 94530, Vol. 3, p 173.
(22) Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry* Wiley & Sons: New York, 1980.

Therefore, no detailed data are reported here for this compound, but the close similarity between uracil and thymine wrt geometry and energetics of nuclear configurations associated with attractive contributions should be pointed out.

Study of the *wave function, electronic density*, and *charge orders* near ring nitrogen nuclei does not reveal recognizable differences in these quantities for pairs of geometric isomers, which differ wrt attractive interactions upon anti→syn conversion. No reason for this surprising result may be given at the present time.

In contrast to electronic density, the *electric field gradients* at ring nitrogen nuclei depend in a systematic and characteristic manner on the number and type of attractive interactions involving such nuclei. This is clearly born out by Table IV, in which the principal values of the field gradient tensor at nitrogen atoms 1 and 3 of the quadruplet of geometric isomers Thy6–9 are collected (HF/6-31G**). These isomers feature the attractive interactions E(aHO·2|sp²·1|aHO·6)(Thy6), E(aHO·2|sp²·1)(Thy7), E(sHO·2|sp²·3) and E(aHO·6|sp²·1)(Thy8), and E(sHO·2|sp²·3)(Thy9).

Consider first the ring site 1 N atom: the field gradient principal components q_{AA} (q_{BB},q_{CC}) increase (decrease) monotonically in the sequence Thy6, Thy8, Thy7, Thy9. In the isomers Thy6, Thy8, and Thy7, respectively, the attractions E(aHO·2|sp²·1|aHO·6)(t_4), E(aHO·6|sp²·1)(t_3), and E(aHO·2|sp²·1)(t_5) are active (all involving sp²·1); in Thy9 none is present. Obviously, the principal values q_{CC} (q_{BB},q_{AA}) in the sequence are directly (oppositely) correlated with t_4, t_3, t_5, and nothing, respectively.

Turning to the field gradient at the ring site 3 N atom, we first note that in the sequence Thy6, Thy8, Thy7, Thy9, attraction E(sHO·2|sp²·3)(t_2) involving this atom occurs with Thy8 and Thy9, but no such interaction is active in Thy6,Thy7. One therefore may expect to find nearly equal values of q_{CC} for isomers Thy8 and Thy9 on one hand and for Thy6 and Thy7 on the other hand. Analogously, one should expect equal values for q_{BB}(q_{AA}) in Thy8,Thy9 and Thy6,Thy7, respectively. For q_{AA} values, the expectation is violated, but these are seen from Table IV to vary only slightly in the sequence Thy6–9. Obviously, the assumption that solely attractive interactions involving sp²·3 should affect the field gradient at the site 3 N atom is too restrictive if finer details are to be understood.

One nevertheless may consider the correlation between attractive interactions involving a nitrogen nucleus and the field gradient acting on it as rather reliable. It will be shown in a forthcoming paper to be essentially correct in the whole set of pyrimidine nucleic acid bases studied so far.

Barriers to Methyl Group Internal Rotation. In the whole set of thymine isomers, eclipsed (ec) CH_3 conformations were found to be more stable than staggered (st) conformations, the latter corresponding to a saddle point. Furthermore, upon ec–st conversion, some of the ISPs experience surprisingly large alterations, but not all of these may be rationalized by repulsive interactions E_{16}(sHO·6|ec(st)CH_3·5) satisfactorily.

The barriers to internal rotation present another complex picture. Assuming the potential to internal rotation to obey the expression (τ = 0 for ec)

$$V(\tau) = {}^1/_2V_3(1 - \cos 3\tau)$$

the barrier V_3 is obtained for each isomer from $V_3 \approx \mathcal{E}_t$(st) – $\mathcal{E}_t$(ec). In table I (last column), the V_3 values are collected as estimated from HF/6-31G** total electronic energies; use of MP2/HF/6-31G** single-point results does not alter the barrier values to any relevant measure. Regarding these data, the following comments apply. (i) Isomers with oxo group in six positions (O:6) feature V_3 values in the range 1.44–1.84 kcal/mol, essentially independent of isomerism of the 2-OH group. (ii) Anti HO·6 isomers nearly always have V_3 values lower by 0.9–1.1 kcal/mol than syn HO·6 isomers, for both 2-hydroxy and 2-oxoisomers; the V_3 values of the 6-anti species ranges between 1 and 1.7 kcal/mol. The latter quantity seems to be higher for isomers with an O:2 substituent. (iii) Thy10 isomers make an exception to both rules, since their V_3 barriers are noticeably lower than those of all other species. No obvious interpretation seems to be available for this finding, but it might be related to the unique topology of the π electron system of the isomers of Thy10.

Final Remarks

The main results of this paper are the following (i) extension of the linear contribution model for conversion energies of thymine geometric isomers and finding of a set of contributions that allow approximate reproduction of conversion energies in the whole set of pyrimidine nucleic acid bases; (ii) determination of correlations between geometric isomerism of OH substituents and CH_3 conformations with ISPs for thymine, which may be qualitatively interpreted in terms of repulsive and attractive interactions and furthermore hold approximately in the whole set; and (iii) assertion of correlations of the electric field gradient at the ring N nuclei with attractive interactions in the set of thymine isomers Thy6–9. Analogous correlations are expected to hold in the whole sets of pyrimidine and purine nucleic bases.

Based on the fact that some of the contribution terms may be estimated to zeroth order from conversion energies of appropriately chosen acyclic compounds, one is tempted to consider these quantities applicable to a larger set of molecules. This holds in the first place for repulsive quantities. For attractive interactions, it appears to be more difficult to speculate on possible realms of validity.

The predictive power of the increment model will be illustrated by the following examples. (i) For the anti→syn conversion

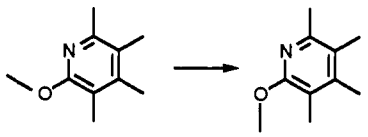
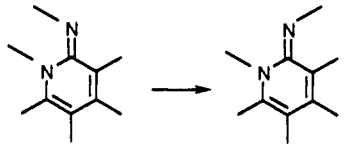


Table VII predicts $\Delta\mathcal{E} \approx -E$(aHO·2|sp²·1) + E_{15}(sHO·2|H·3) ≈ –(–4.8(5))+1.7(3) ≈ +6.5(8) kcal/mol; at the HF/6-31G** level, $\Delta\mathcal{E}$ ≈ +5.99 kcal/mol is found. (ii) For the conversion



one would predict $\Delta\mathcal{E} \approx -E_{15}$(aHN:6|H1) + E(H·1|sp²N:6) + E_{15}(sHN:6|H·5) ≈ –(+0.6 to +1.2) + (–3.5) + (+2.0) ≈ –2.1 to –2.7 kcal/mol. At the HF/6-31G** level, the value is $\Delta\mathcal{E}$ ≈ –2.81 kcal/mol. As mentioned above, the first increment E_{15}-(aHN:6|H·1) is not well determined and possibly may be improved by inclusion of this conversion in the analysis.

The foregoing investigations are based on HF/6-31G** predictions. If MP2/HF/6-31G** values with or without inclusion of zero-point energy contributions are to be used, the contributions are reduced in magnitude by 5–8%. The repulsive interaction E_{15}(aHN:6|H·1) makes an exception to this rule; however, it has not yet been possible to obtain a better determination of this term.

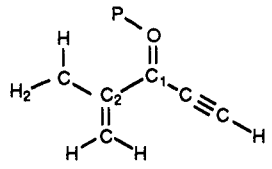
As far as quantum chemical energy and structural data at the HF/6-31G** and higher levels are available in the literature for cytosine, uracil, and thymine, full agreement with the present work has been noted. It recently has been shown for the three lowest isomers U1, U3, and U8 of uracil by Leszczynki[23] that

(23) Leszynski, J. *J. Phys. Chem.* **1992**, *96*, 1649.

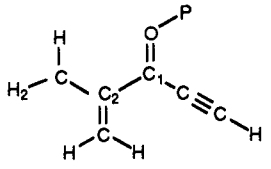
higher approximations than those of MP2/6-31G** or configuration interaction approaches do not yield a monotonic dependence of conversion energies on the level of approximation.[23] Most of the increments to conversion energies of geometrical isomers reported here may, however, be expected to remain unaffected within the present error limits.

Appendix

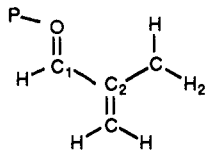
Model Compounds for Thymine. The following model compounds have been treated by the same quantum chemical approximation as for thymine (p denotes proton, charge number $z = +1$, ec and st denote CH_3 conformations eclipsing and staggering the C=O bond, respectively):



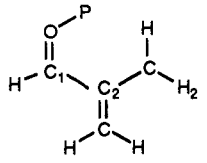
MThy1 syn (ec, st)



MThy1 anti (ec, st)



pOC8 anti (ec)



pOC8 syn (ec)

O-Protonated carbonyl compounds were earlier[2,3] found to have similar geometric isomer conversion energies as imines and were also used to obtain zeroth-order estimates for conversion of geometric isomers of enolic OH groups, though the latter procedure has not yet been fully justified. Total electronic energies are collected in Table XI. In its right side, conversions energies

Table XI. Electronic Energy Data for Thymine Model Compounds

compound	$-\mathcal{E}_t$ (au)[a]		$\Delta\mathcal{E}_t$ (kcal/mol)	
MThyls ec	305.820 70	0	0	
MThyla ec	305.824 27	−2.24		0
MThyls st	305.816 91	0	2.38	
MThyla st	305.821 79	−3.06		1.56
pOC8a st	230.145 77	0		
pOC8s st	230.143 80	1.24		

[a] HF/6-31G** data.

are listed, which may be related to contributions as follows (kcal/mol, a and s stand for anti and syn)

$$\Delta\mathcal{E}_{s\to a} = -2.24(-3.06)$$

$$\approx -E_{16}^{aux}(pO{:}1|ecCH_3{\cdot}2\|2{:}CH_2) \qquad (1\text{-A1})$$

$$\approx -E_{16}^{aux}(pO{:}1|stCH_3{\cdot}2\|2{:}CH_2) \qquad (1\text{-Al st})$$

$$\Delta\mathcal{E}_{8a\to s} = +1.24$$

$$\approx -E_{14}^{aux}(apO{:}2|H{\cdot}1) + E_{16}^{aux}(spO{:}2|ecCH_3{\cdot}2\|2{:}CH_2) \qquad (2\text{-A2 ec})$$

$E_{14}(spO{\cdot}2|H{\cdot}1)$ was estimated earlier to be 1.00(15),[2] yielding from (2-A2ec) a second estimate for $E_{16}^{aux}(pO{:}1|ecCH_3{\cdot}2\|2{:}CH_2) \approx 2.24$.

Since neither of the two zeroth-order estimates for this quantity can be given any perference, both should be used with equal weight in the least-squares process.

Acknowledgment. The authors express their gratitude for large grants of free computer time by the ETH Zuerich administration and Sandoz, Inc., Basle, for support of this work. Typing of the manuscript by Mrs. K. Gunthard is gratefully acknowledged. Also, we wish to thank Dr. B. Schweizer for help with data files.