The Tautomeric Equilibria of Thio Analogues of Nucleic Acid Bases. Part 2.[1] AM1 and *ab initio* Calculations of 2-Thiouracil and its Methyl Derivatives

Alan R. Katritzky* and Miroslaw Szafran‡
Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA
John Stevens
Corporate Research Laboratories/3M, St. Paul, Minnesota 55144, USA

Relative stabilities predicted by the AM1 method, and relative energies from 3-21G* and 6-31G** calculations agree with qualitative and quantitative experimental vapour phase data for 2-thiouracil and for its four tautomeric monomethyl derivatives. The AM1 proton affinities for the more stable tautomeric forms are usually significantly lower than, but almost within experimental values of, the measured PA values. A plot of experimental *vs.* calculated PA gives a straight line of almost unit slope.

2-Thiouracil has been identified in t-RNA[2] and is known to possess important biological activity, *e.g.* mutagenic and anticancer.[3] In spite of its biochemical interest, very few systematic theoretical[4] or experimental[5] studies are available of the tautomerism of 2-thiouracil. Recently, the proton affinities (PA) of 2-thiouracil and of all four of its monomethyl and of four of its six dimethyl derivatives, have been measured.[1] The structures of the most stable tautomers of thiouracil and of its four monomethyl derivatives were recently determined from i.r. spectra in the vapour phase and in inert low temperature matrices.[5]

In the present paper, the geometries, dipole moments, and relative energies have been calculated, using the AM1 method, for the most stable conformer of each of the individual tautomers of 2-thiouracil, and of all the tautomers of all the methyl and dimethyl derivatives and for all the tautomers of all the mono protonated cations of all these species (see Figure). The results are compared with i.r. data and with measured PA values. For some of these species, *ab initio* 3-21G* and 6-31G**, results are also presented.

Calculations.—All calculations were carried out with complete geometry optimisation. The MNDO geometries of the 2-thiouracil tautomers[4b] were taken as the starting points. Planarity of the 2-thiouracil ring was assumed, but the conformations of OH, SH, OMe, and SMe substituent groups were optimised (only data for the most stable conformers are described in this paper). The AM1[6] and MNDO[7] calculations were performed using the MOPAC program[8] on a MicroVAX II. The *ab initio* calculations at the 3-21G* level were carried out using the modified MNDO method applied in the *ab initio* suite of programs from QUIPU.[9] The restricted Hartree–Fock method was used with the basis sets defined in the reports, and the search for stable geometry was ended when the component of the gradients reached 0.001. The single-point calculations in the 6-31G** basis set were performed with program GAUSSIAN 86.[10]

Results and Discussion

2-Thiouracil can exist in six tautomeric forms with unbroken cyclic conjugation (aromatic tautomers) as shown in the Figure (non-ring conjugated, *i.e.* non-aromatic, tautomers also exist, but are likely to be of little importance[11]). Additionally, some tautomers can exist as more than one conformer. In this paper we consider the most stable conformer corresponding to each tautomeric form.

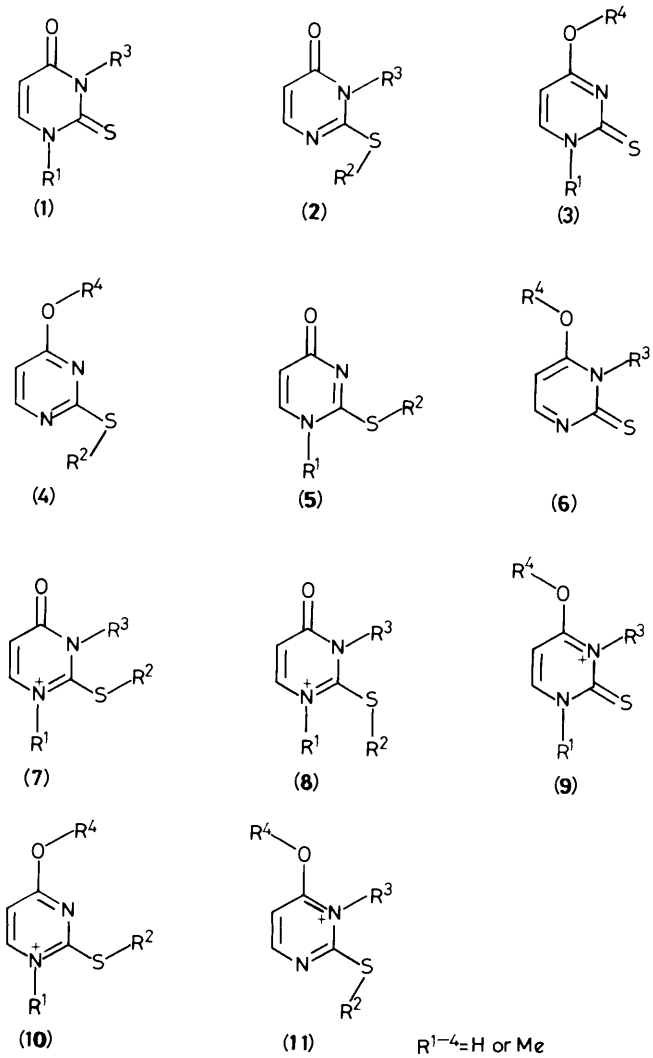

Figure.

‡ On leave from Dept. of Chemistry, A. Mickiewicz University, 60780 Poznan, Poland.

Table 1. Optimised bond lengths/Å and bond angles/° of three 2-thiouracil tautomers and two methyl derivatives.

	(1) ($R^1 = R^3 = H$)			(2) ($R^2 = R^3 = H$)		(4) ($R^2 = R^4 = H$)		(2) (R^2 = Me, $R^3 = H$)		(4) (R^2 = Me, $R^4 = H$)	
Parameter	3-21G*	AM1	Exp[a]	3-21G*	AM1	3-21G*	AM1	3-21G*	AM1	3-21G*	AM1
1-2	1.361	1.400	1.359	1.280	1.333	1.319	1.369	1.283	1.333	1.319	1.366
2-3	1.355	1.387	1.368	1.359	1.381	1.334	1.363	1.361	1.384	1.338	1.367
3-4	1.404	1.415	1.380	1.404	1.423	1.314	1.364	1.404	1.419	1.311	1.363
4-5	1.455	1.468	1.417	1.450	1.463	1.392	1.419	1.451	1.462	1.394	1.419
5-6	1.326	1.359	1.360	1.336	1.368	1.369	1.402	1.338	1.367	1.368	1.402
6-1	1.380	1.384	1.368	1.389	1.381	1.344	1.350	1.386	1.382	1.345	1.350
2-5	1.661	1.590	1.662	1.759	1.715	1.752	1.710	1.754	1.719	1.747	1.714
4-O	1.210	1.240	1.236	1.214	1.241	1.344	1.368	1.214	1.242	1.344	1.369
5-H	1.066	1.097		1.067	1.097	1.066	1.094	1.067	1.096	1.066	1.094
6-H	1.069	1.106		1.069	1.105	1.069	1.105	1.069	1.105	1.069	1.106
N_1-H(C)	0.999	0.997									
N_3-H(C)	1.002	1.000		1.002	0.997			1.002	0.997		
S_2-H(C)			1.323	1.335	1.324	1.334	1.812	1.734	1.810	1.731	
O_4-H(C)						0.968	0.973			0.968	0.973
1-2-3	114.2	118.0	114.6	123.0	125.4	124.4	126.6	122.5	124.9	124.1	126.7
2-3-4	127.8	123.1	125.5	123.6	119.6	118.2	115.5	124.2	119.9	118.4	114.9
3-4-5	113.4	116.0	116.2	112.2	115.1	121.9	122.6	112.2	115.3	122.0	123.3
4-5-6	119.7	120.1	119.7	120.3	119.4	115.9	116.2	119.7	119.7	115.8	115.7
5-6-1	121.8	121.4	119.5	123.5	123.4	122.5	123.1	124.3	123.5	122.5	123.1
6-1-2	123.2	121.3	124.6	117.3	116.5	117.1	116.0	117.1	116.7	117.2	116.2
R-1-H(C)	116.3	117.4									
1-2-S	123.3	120.0	123.1	121.2	119.5	120.0	117.7	122.4	120.2	120.8	118.8
2-3-H(C)	116.5	117.6		121.0	121.0			115.8	120.7		
3-4-O	120.3	118.8	117.6	120.3	118.2	118.5	120.3	120.3	118.3	118.5	120.0
1-6-H(C)	115.4	115.7		114.7	114.1	116.3	115.9	120.0	114.1	121.2	115.7
6-5-H	122.3	123.0		122.4	123.5	123.3	122.8	120.0	123.6	120.9	123.1
2-3-H(C)				94.3	99.6	95.2	99.2	99.6	105.2	100.7	105.3
4-O-H(C)						111.1	109.7			111.0	109.8

[a] Data for 6-methyl-2-thiouracil (C. Delage, A. H'Naifi, M. Goursolle, A. Carpy, *C. R. Seances Acad. Sci. Paris*, Serie II, 1986, **302**, 219).

For two of the tautomeric forms of 2-thiouracil [(**3**) and (**6**)], the conformers shown in structures (**3**) and (**6**) are more stable than those proposed by Buda:[4b] the differences in the heats of formation between the conformers described in this paper and those for the conformers described by Buda [4b] are 0.3 and 1.6 kcal mol^{-1} for (**3**) and (**6**), respectively.

Conformations of methyl derivatives of 2-thiouracil can be defined by their dihedral angle. These angles for the most stable conformers are as follows:

N_1-Me ∢ H–C–N_1–C_6 = 0	(**1**), (**3**), (**5**), (**7**)–(**10**)
N_3-Me ∢ H–C–N_3–C_4 = 0	(**1**), (**2**), (**6**), (**7**)–(**9**), (**11**)
S-Me ∢ H–C–S–C_2 = 180	(**2**), (**4**), (**5**), (**7**), (**8**), (**10**), (**11**)
O-Me ∢ H–C–O–C_4 = 180	(**3**), (**4**), (**6**), (**9**), (**10**), (**11**)

Optimised Geometry.—Optimised bond lengths and angles are given in Table 1 for three of the 2-thiouracil tautomers and for two of the methyl derivatives studied. For the carbonyl-thione tautomer of 2-thiouracil (**1**), the result of the 3-21G* basis set calculation agrees with the experimental *X*-ray data to within 0.04 Å and 2.8° for bond lengths and angles, respectively. Comparisons of the AM1 geometries of these 5 compounds with those from the 3-21G* set shows that: (*i*) the ring bond lengths obtained by the AM1 method are, on average, longer by 0.026 Å (sample standard deviation σ_{n-1} = 0.017 Å) than the bond lengths determined by the 3-21G* basis set; (*ii*) C–H and C–O bonds predicted by the AM1 method are longer by 0.033 Å (σ_{n-1} = 0.004) and 0.027 Å (σ_{n-1} = 0.002) respectively, by comparison with those optimised by the 3-21G* basis set; (*iii*) the 2C–S bonds calculated by the AM1 method are shorter than those by the 3-21G* basis set by 0.048 Å (σ_{n-1} = 0.016 Å); and (*iv*) differences in the ring bond angles between the AM1 and 3-21G* data are within ± 3°.

Relative Stabilities.—The calculated values for the heats of formation and for the dipole moments of 2-thiouracil, of its mono- and dimethyl derivatives, and of their protonated analogues, are collected in Tables 2–4.

For each of the potentially tautomeric compounds the relative stabilities predicted by the AM1 method (Table 2), and relative energies predicted by the 3-21G* and 6-31G** basis sets (Table 4), agree well with the results of the i.r. spectral determinations of the dominant tautomer in the vapour phase.[5] The 3-21G* calculations predict larger differences between the individual tautomeric forms in comparison with the AM1 method. In the case of 2-methylthiouracil, two tautomers (**2**) and (**4**) were observed in the vapour phase i.r. spectrum. Our calculations predict that tautomer (**2**) (R^2 = Me; R^3 = H) is more stable by 2.95 kcal mol^{-1} using the AM1 method (Table 2), or by 3.96 kcal mol^{-1} (3-21G*), or by 1.64 kcal mol^{-1} (6-31G**) (Table 4) than tautomer (**4**) (R^2 = Me; R^4 = H). In this particular example, it seems that the 6-31G** calculations give the best relative energies for the investigated tautomer pair.

The relative stabilities predicted by the MNDO method (Table 2) disagree with the i.r. spectral reults. The introduction of corrections proposed by Dewar and McKee [12] improves the correlation with the i.r. spectra,[4b] however, the AM1 method is clearly superior to MNDO.

Heats of Formation and Proton Affinities.—The absolute values of the heats of formation and of the proton affinities calculated using the AM1 method for the most stable tautomers are listed in Table 5. In most cases, the calculated proton affinities are lower than the experimental data, but the differences are comparable with the experimental errors.[13,14] A plot of experimental PA *versus* calculated PA gives a straight line (PA_{exp} = 15.72 ± 0.95 PA_{calc}, R = 0.94). Similar correl-

Table 2. Calculated values of heats of formation (ΔH_f/kcal mol^{-1}), relative stability (RS/kcal mol^{-1}), and dipole moments (D) of 2-thiouracil and its mono- and di-methyl derivatives.

Structural type					Heat of formation				Relative stability			Dipole moment		Tautomer observed
No.	R^1	R^2	R^3	R^4	AM1	Est.[d]	MNDO	MNDO[c]	AM1	MNDO	MNDO[c]	AM1	MNDO	in i.r.[a]
2-Thiouracil tautomers														
(1)	H	—	H	—	−1.14		−13.94	−14.77	0	3.91	3.75	4.92	4.94	Yes
(2)	—	H	H	—	5.08		−9.21	−5.85	6.22	8.64	12.67	2.84	2.46	No
(3)	H	—	—	H	16.39		−12.15	−13.51	17.53	5.1	5.01	8.39	6.16	No
(4)	—	H	—	H	7.81		−17.85	−18.52	8.95	0	0	1.60	2.10	No
(5)	H	H	—	—	15.81		−0.64	−1.41	16.95	17.21	17.11	5.65	5.52	No
(6)	—	—	H	H	12.81		−5.29	−6.04	13.94	12.56	12.48	7.35	6.24	No
1-Methyl-2-thiouracil tautomers														
(5)	Me	H	—	—	22.31	22.4	3.29		16.62	13.00		6.15	5.80	No
(1)	Me	—	H	—	5.69	5.5	−9.71		0	0		5.26	5.25	Yes
(3)	Me	—	—	H	23.02	23.0	−2.46		17.33	7.25		8.37	8.47	No
2-Methyl-2-thiouracil tautomers														
(5)	H	Me	—	—	11.43	11.6	−9.14		10.64	16.61		5.51	5.18	No
(2)	—	Me	H	—	0.79	0.9	−17.53		0	8.22		3.51	3.19	Yes
(4)	—	Me	—	H	3.74	3.6	−25.75		2.95	0		1.43	1.26	Yes
3-Methyl-2-thiouracil tautomers														
(1)	H	—	Me	—	6.51	6.4	−7.18		0	0		4.13	4.17	Yes
(2)	—	H	Me	—	12.32	12.6	−2.70		5.81	4.48		2.52	2.17	No
(6)	—	—	Me	H	20.51	20.3	−0.08		14.00	7.10		7.27	7.51	No
4-Methyl-2-thiouracil tautomers														
(3)	—	—	H	Me	22.81	22.4	2.42		9.20	11.17		8.70	8.53	No
(4)	—	H	—	Me	13.61	13.8	−9.75		0	0		1.46	1.42	Yes[b]
(6)	—	—	H	Me	18.50	18.8	0.18		4.89	9.93		8.18	8.29	No
Dimethyl-2-thiouracil isomers														
(5)	Me	Me	—	—	18.12	18.2	−5.04					6.00	5.45	—
(1)	Me	—	Me	—	13.76	13.0	−2.10					4.48	4.44	—
(2)	—	Me	Me	—	8.10	8.4	−10.77					3.19	2.90	—
(4)	—	Me	—	Me	9.61	9.6	−17.20					1.05	1.11	—
(6)	—	—	Me	Me	27.19	26.3	8.61					7.97	8.03	—
(3)	Me	—	—	Me	29.48	29.0	6.08					8.62	8.52	—

[a] From ref. 5. [b] $R^4 = CH_2C(CH_3)_3$. [c] From ref. 36. [d] Estimated using parameters listed in Table 6 (see text).

Table 3. Calculated values of heats of formation (kcal mol^{-1}) and dipole moments (D) of protonated 2-thiouracil and protonated mono- and di-methyl thiouracils.

Structural type							
No.	R^1	R^2	R^3	R^4	ΔH_f	$\Delta H_{f(est)}$[a]	RS
Protonated 2-thiouracil tautomers							
(7)	H	H	H	—	167.70	—	5.18
(8)	H	H	H	—	168.55	—	6.03
(9)	H	—	H	H	165.56	—	3.04
(10)	H	H	—	H	162.52	—	0
(11)	—	H	H	H	165.69	—	3.17
Protonated 1-methyl-2-thiouracil tautomers							
(7)	Me	H	H	—	171.70	171.5	5.7
(10)	Me	H	—	H	166.00	166.3	0
(9)	Me	—	H	H	169.00	169.4	3.6
Protonated 2-methyl-2-thiouracil tautomers							
(7)	H	Me	H	—	158.90	160.1	3.31
(11)	—	Me	H	H	158.62	158.1	3.03
(10)	H	Me	—	H	155.54	154.9	0
Protonated 3-methyl-2-thiouracil tautomers							
(7)	H	H	Me	—	171.96	172.1	2.17
(9)	H	—	Me	—	170.36	170.0	0.57
(11)	—	H	Me	H	169.79	170.1	0
Protonated 4-methyl-2-thiouracil tautomers							
(10)	H	H	—	Me	165.35	165.2	0
(11)	—	H	H	Me	168.36	168.4	3.01
(9)	H	—	H	Me	168.11	168.3	2.76

Structural type							
No.	R^1	R^2	R^3	R^4	ΔH_f	$\Delta H_{f(est)}$[a]	RS
Protonated 1,2-dimethyl-2-thiouracil tautomers							
(7)	Me	Me	H	—	163.25	163.9	3.9
(10)	Me	Me	—	H	159.35	158.7	1
Protonated 1,3-dimethyl-2-thiouracil tautomers							
(7)	Me	H	Me	—	176.65	175.9	1.72
(9)	Me	—	Me	H	174.93	173.8	0
Protonated 2,3-dimethyl-2-thiouracil tautomers							
(8)	H	Me	Me	—	164.59	165.3	1.51
(11)	—	Me	Me	H	163.08	162.5	0
Protonated 2,4-dimethyl-2-thiouracil tautomers							
(10)	H	Me	—	Me	158.71	157.6	0
(11)	—	Me	H	Me	161.49	160.8	2.78
Protonated 3,4-dimethyl-2-thiouracil tautomers							
(9)	H	—	Me	Me	173.42	172.7	0.48
(11)	—	H	Me	Me	172.94	172.8	0
Protonated 1,4-dimethyl-2-thiouracil tautomers							
(10)	Me	H	—	Me	169.13	169.0	0
(9)	Me	—	H	Me	172.34	172.1	3.21

[a] Estimated using parameters listed in Table 6 (see text).

Table 4. *ab initio* total energies (Hartrees)[a] and relative energies (kcal mol^{-1}) of some 2-thiouracil tautomers.

Structural type					Total energies		Relative energies	
No.	R^1	R^2	R^3	R^4	3-21G*	6-31G**[b]	3-21G*	6-31G**
2-Thiouracil tautomers								
(1)	H	—	H	—	−731.375 126 3	—	0	—
(2)	—	H	H	H	−731.339 075 6	—	22.6	—
(4)	—	H	—	H	−731.334 167 6	—	25.7	—
2-Methyl-2-thiouracil tautomers								
(2)	—	Me	H	—	−770.162 524 7	−774.117 227 1	0	0
(4)	—	Me	—	H	−770.156 213 9	−774.114 617 9	3.96	1.64

[a] 1 a.u. or Hartree = 627.5 kcal mol^{-1}. [b] All calculations were carried out with the optimised 3-21G* geometries.

Table 5. Experimental and calculated proton affinities for conversion of dominant free-base tautomers into the most stable mono cations.

Free bases (B)						Protonated cation (B^+H)						PA[a]	
Structural type					ΔH_f	Structural type					ΔH_f	Calc.	Exp.
No.	R^1	R^2	R^3	R^4		No.	R^1	R^2	R^3	R^4			
(1)	H	—	H	—	−1.14	(10)	H	H	—	H	162.57	203.5	209.1 ± 0.4
(1)	Me	—	H	—	5.69	(10)	Me	H	—	H	166.00	206.9	214.1 ± 0.6
(2)	—	Me	H	—	0.79	(10)	H	Me	—	H	155.59	212.4	220.7 ± 0.1
(1)	H	—	Me	—	6.51	(11)	—	H	Me	H	169.79	203.9	209.1 ± 0.3
(4)	—	H	—	Me	13.61	(10)	H	H	—	Me	165.35	215.5	215.2 ± 0.3
(5)	Me	Me	—	—	18.12	(10)	Me	Me	—	H	159.35	226.0	233.2 ± 0.6
(1)	Me	—	Me	—	13.76	(9)	Me	—	Me	H	174.93	206.0	214.1 ± 0.6
(2)	—	Me	Me	—	8.10	(11)	—	Me	Me	H	163.08	212.2	217.5 ± 0.4
(4)	—	Me	—	Me	9.61	(10)	H	Me	—	Me	158.71	218.1	223.0 ± 1.2
(6)	—	—	Me	Me	27.19	(11)	—	H	Me	Me	172.94	221.4	
(3)	Me	—	—	Me	29.48	(10)	Me	H	—	Me	169.13	227.5	

[a] PA = 367.2 + ΔH_f(B) − ΔH_f(B^+H); (from M. J. S. Dewar and K. M. Dieter, *J. Am. Chem. Soc.*, 1986, **108**, 8075).

Table 6. Incremental effects of methyl substitution on heats of formation (kcal mol^{-1}) in 2-thiouracil and its protonated analogues.

	B	B^+H
N_1–CH_3	6.6	3.8
N_3–CH_3	7.5	4.4
S_2–CH_3	−4.2	−7.6
O_4–CH_3	6.0	2.7

ations have been found between PA_{exp} and proton affinity calculated using *ab initio* methods for simpler compounds.[14]

Effects of Methyl Groups on Heats of Formation.— Substitution of N–H, O–H, and S–H protons by a methyl group caused systematic changes in heats of formation (Table 6). Methyl groups attached to nitrogen and oxygen atoms increase ΔH_f. However, at an S atom, a methyl group decreases the ΔH_f value. Methyl groups cause different effects on neutral and protonated compounds. The corrections listed in Table 6, when added to the heats of formation of the corresponding tautomers of 2-thiouracil (or its protonated analogues), reproduce the calculated heats of formation of the mono- and dimethyl derivatives with an average accuracy of ±0.15 kcal mol^{-1} (see Table 2).

Conclusions

The results of these calculations, and comparison with the experimental results of the preceding paper suggest that the AM1 method is of real value for the calculation of heats of formation of individual tautomers, and for the calculation of proton affinities.

Acknowledgements

This work was partly supported by the Polish Academy of Sciences (CPBP 01.12.8.5). We thank a referee for helpful comments and Professor M. J. S. Dewar for communication of unpublished work.

References

1 Part 1, A. R. Katritzky, G. Baykut, S. Rachwal, M. Szafran, K. C. Caster, and J. Eyler, 1989, 1491.
2 (*a*) J. Carbon, H. David, and M. H. Studier, *Science*, 1968, **161**, 1146; (*b*) L. Baczynskyj, K. Biemann, and R. H. Hall, *ibid.*, **159**, 1481.
3 (*a*) E. Galkiewicz, M. Pyziak, J. Chomiczewski, and T. Gorski, *Med. Dosw. Mikrobiol.*, 1979, **31**, 11; (*b*) E. B. Astwood, A. Bessll, and A. M. Hugnes, *Endocrinology*, 1945, **37**, 456.
4 (*a*) M. Geller, A. Pohorille, and A. Jaworski, *Biochim. Biophys. Acta*, 1973, **331**, 1; (*b*) A. B. Buda, *J. Mol. Struct.(Theochem.)*, 1987, **149**, 185.
5 (*a*) A. Psoda and D. Shugar, *Acta Biochem. Pol.*, 1979, **26**, 55, and ref. therein; (*b*) H. Rostkowska, A. Barski, K. Szczepaniak, M. Szczesniak, and W. B. Person, *J. Mol. Struct.*, 1988, **176**, 137, and ref. therein.
6 M. J. S. Dewar, E. G. Zoebish, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902.
7 M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, 1977, **99**, 4899.
8 J. J. P. Stewart, MOPAC Program Package, *QCPE*, 1983, No. 455 (version 3.0), AM1 parameters for 5 were provided by Professor Dewar.
9 G. D. Puruis III, 'QUIPU V 3.6—A Molecular Modelling System,' Quantum Theory Project, Technical Report TF 1020, University of Florida, Gainesville, Florida 32611.
10 M. J. Frisch, J. S. Binkley, H. B. Schlagel, R. Raghavacheri, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowitcz, C. M. Rohlfing, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whiteside, D. J. Fox, E. M. Fleuder, and J. A. Pople, Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburg PA, 1984.

11 J. Elguero, C. Marzin, A. R. Katritzky, and P. Linda, 'The Tautomerism of Heterocycles,' eds. A. R. Katritzky and A. J. Boulton, Academic Press, London, 1976.
12 M. J. S. Dewar and M. L. McKee, *J. Comput. Chem.*, 1983, **4**, 84.
13 S. G. Lias, J. F. Liebman, and R. D. Levin, *J. Chem. Phys. Ref. Data*, 1984, **13**, 695.
14 I. A. Koppel, U. H. Molder, and V. A. Palm, *Organic Reactivity* (*Tartu*), 1985, **22**, 3.

Received 13*th November* 1988; *Paper* 8/04549F