Structure and Tautomerism of the Neutral and Monoanionic Forms of 4-Thiouracil Derivatives

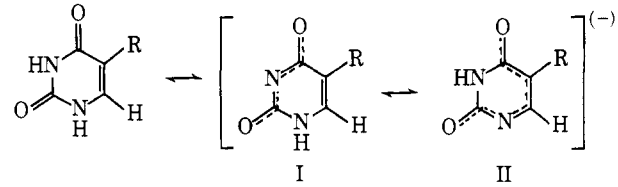
Anna Psoda, Z. Kazimierczuk, and D. Shugar*

Contribution from the Department of Biophysics, Institute of Experimental Physics, University of Warsaw, 02-089 Warsaw, Poland. *Received April 3, 1974*

Abstract: A study of the ultraviolet and infrared spectra of a variety of 4-thiouracil derivatives in various solvents demonstrated that 4-thiouracil, 4-thiouridine, and 5,6-dihydro-4-thiouracil exist in the 2-keto-4-thione form both in aqueous and nonaqueous media. 4-Methylthiouracil is in the 1(*H*)-2-keto form and undergoes protonation on the ring at N_3. Ionization of 4-thiouracil leads to formation of an equilibrium mixture of two monoanions, one of which results from dissociation of the N_1 proton, the other from dissociation of the N_3 proton. Each monoanion involves extensive charge delocalization between O^2 and S^4, *via* the ring N_3 when the ring N_3 proton dissociates, and *via* the 5,6-bond when the N_1 proton dissociates. The ratio of the two tautomeric monoanions was experimentally determined with the aid of the first derivative uv spectrum of 4-thiouracil monoanion and independently from the infrared spectra of the monoanions of 4-thiouracil and its two N-monomethylated derivatives, with results in reasonably good agreement. The ratio of the two monoanions in ionized 4-thiouracil may be markedly shifted by variation of the solvent medium, as previously demonstrated for tautomeric monoanion mixtures of 2,4-dioxopyrimidines. Some tentative infrared band assignments are presented.

Of the four thiopyrimidines hitherto identified as minor bases in tRNA, 4-thiouracil and its nucleoside, 4-thiouridine, have been subjected to the most intensive investigation.[1] It is therefore surprising that no systematic study has yet been devoted to the possible tautomeric forms of the neutral and ionic species of 4-thiouracil and its derivatives, notwithstanding that all the relevant reference compounds required for this purpose have already been reported. Some evidence, based on comparisons of uv absorption spectra, has been adduced for the 2-keto-4-thione structure of the neutral forms in aqueous medium of 4-thiouracil,[2,3] although Scheit[4] was unable to detect any signal for a ring N_3 proton in the nmr spectrum of a 4-thiouridine analog in DMSO. X-Ray diffraction shows that 4-thiouracil nucleosides possess the 2-keto-4-thione form in the solid state.[5] Practically no attention has been drawn to the structures of the ionic forms, although these normally exist in appreciable proportions at physiological pH because of the much lower p*K* of 4-thiouracil as compared to uracil (see Table I, below).

In the present investigation we have examined the structures of the neutral and ionic forms of various 4-thiouracil derivatives in the light of previous findings on 2,4-diketopyrimidines, for which the monoanionic species consist of an equilibrium mixture of two



I II (1)

R = H, alkyl, halogen

anions[6-8] each with extensive charge delocalization,[8] as illustrated by eq 1. Both uv and ir spectroscopy, in aqueous and nonaqueous media, were employed insofar as solubilities of the appropriate model compounds permitted.

Experimental Section

Materials. 4-Thiouracil was prepared according to Ikehara, *et al.*,[9] 1-methyl-4-thiouracil as described by Fox, *et al.*,[10] 3-methyl-4-thiouracil according to Ueda and Fox,[2] 1,3-dimethyl-4-thiouracil according to Elion and Hitchings,[11] 2-oxo-4-methylthiopyrimidine as described by Wheeler and Johnson,[12] and 1-methyl-2-oxo-4-methylthiopyrimidine by the procedure of Ikehara, *et al.*[9] Thiation of uridine in dioxane[13] yielded 4-thiouridine in 72% yield, mp 136–138°, as compared to 139–140° reported by Scheit.[4] Similar thiation of 1-cyclohexyluracil[14] gave 1-cyclohexyl-4-thiouracil, mp 202–203°.[14] We are indebted to Dr. Vinko Skaric for a gift of 1-methyl-5,6-dihydro-4-thiouracil[13] and to J. Giziewicz for the syntheses of 2,2′-anhydro-1-(β-D-arabinofuranosyl)uracil (2,2′-anhydro-

(1) See, *e.g.*, E. Sato and Y. Kanaoka, *Tetrahedron Lett.*, 40, 3547 (1969); T. Sarmeyima, M. Kita, M. Sameyoshi, and F. Sawada, *Biochim. Biophys. Acta*, **179**, 1 (1969); R. H. Hall, "The Modified Nucleosides in Nucleic Acids," Columbia University Press, New York, N. Y., 1971; L. Schulman, *J. Mol. Biol.*, **58**, 117 (1971); K. H. Scheit, *Biochim. Biophys. Acta*, **209**, 445 (1971); M. Yaniv, A. Chestier, F. Gros, and A. Favre, *J. Mol. Biol.*, **58**, 381 (1971); N. J. Leonard, D. E. Bergstrom, and G. L. Tolman, *Biochem. Biophys. Res. Commun.*, 44, 1524 (1972); D. E. Bergstrom and N. J. Leonard, *J. Amer. Chem. Soc.*, **94**, 6178 (1972); *Biochemistry*, **11**, 1 (1972); K. Miura, M. Shiga, and T. Ueda, *J. Biochem. (Tokyo)*, **73**, 1279 (1973); H. Iwamura, *Biochim. Biophys. Acta*, **308**, 333 (1973); M. Wrona and B. Czochralska, *Electroanal. Chem.*, **48**, 43 (1973); A. Favre, *Photochem. Photobiol.*, **19**,1 5 (1974), and other pertinent references cited in the foregoing.

(2) T. Ueda and J. J. Fox, *J. Amer. Chem. Soc.*, **85**, 4024 (1963).

(3) H. Maehr, M. Leach, and V. Toome, *J. Heterocycl. Chem.*, **9**, 1390 (1972).

(4) K. H. Scheit, *Tetrahedron Lett.*, 113 (1967).

(5) W. Saenger, *J. Amer. Chem. Soc.*, **94**, 621 (1972); W. Saenger and D. Suck, *Eur. J. Biochem.*, 32, 473 (1973); W. Saenger and K. H. Scheit, *J. Mol. Biol.*, **50**, 153 (1970).

(6) K. Nakanishi, N. Suzuki, and P. Yamazaki, *Bull. Chem. Soc. Jap.*, 34, 53 (1961).

(7) I. Wempen and J. J. Fox, *J. Amer. Chem. Soc.*, **86**, 2474 (1964).

(8) K. L. Wierzchowski, E. Litońska, and D. Shugar, *J. Amer. Chem. Soc.*, **87**, 4621 (1965).

(9) M. Ikehara, T. Ueda, and K. Ikeda, *Chem. Pharm. Bull.*, 10, 767 (1962).

(10) J. J. Fox, D. V. Praag, I. Wempen, I. Doerr, L. Cheong, J. E. Knoll, M. L. Edinoff, A. Bendich, and G. B. Brown, *J. Amer. Chem. Soc.*, **81**, 178 (1959).

(11) G. B. Elion and G. H. Hitchings, *J. Amer. Chem. Soc.*, **69**, 2138 (1947).

(12) W. L. Wheeler and T. B. Johnson, *Amer. Chem. J.*, 42, 30 (1909).

(13) V. Skaric, B. Gaspert, I. Jerkunica, and D. Skaric, *Croat. Chem. Acta*, **37**, 199 (1965); E. A. Falco, B. A. Otter, and J. J. Fox, *J. Org. Chem.*, **35**, 2326 (1970); R. S. Klein, I. Wempen, K. A. Watanabe, and J. J. Fox, *ibid.*, **35**, 2330 (1970); T. Kulikowski and D. Shugar, *J. Med. Chem.*, **17**, 269 (1974).

(14) J. T. Kuśmierek and D. Shugar, *Acta Biochim. Polon.*, **17**, 259 (1970).

araU), as described by Hampton and Nichol,[15] and 2,2'-anhydro-1-(β-D-arabinofuranosyl)-4-thiouracil (2,2'-anhydro-4-thio-araU) according to the same procedure with 4-thiouridine as starting product. The melting point and uv spectrum of the 2,2'-anhydro-4-thio-araU were identical with those reported by Doerr and Fox[16] for the same compound obtained by an alternative route. All compounds were chromatographically homogeneous in several solvent systems.

Methods. Uv absorption spectra were run on a Zeiss (Jena, DDR) VSU-2 spectrophotometer with 10-mm pathlength matched cuvettes. A Radiometer PHM-4d compensating pH meter with glass electrode was employed for pH measurements. Acetate and phosphate buffers, 0.01 *M*, were routinely employed. Extremes of pH were obtained with 0.01 *N* HCl (pH 2), 0.1 *N* HCl (pH 1), 0.01 *N* NaOH (pH 12), 0.1 *N* NaOH (pH 13), and 1.0 *N* NaOH (pH 14). Spectral titrations to establish the neutral and ionic forms, as well as the associated pK_a values, were performed as elsewhere described.[17] Although pH measurements were carried out to an accuracy of 0.01 pH units, the errors involved in the various manipulations limited the accuracy of the measured pK values to ±0.05 pH units.

First derivative spectra were recorded directly on an instrument constructed as described by Balslev[18] but with the modulated slit replaced by a modulated mirror system.

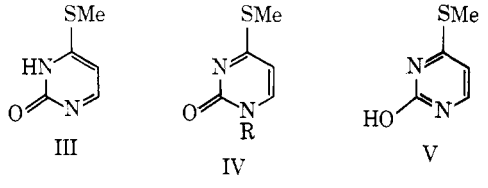
Infrared absorption spectra were recorded on a Zeiss (Jena, DDR) double-beam UR-10 spectrophotometer, using variable pathlength cuvettes fitted with CaF_2 windows. Solvents employed included Eastman-Kodak DMSO, which was dried over CaH_2, Koch-Light D_2O, the deuterium content of which exceeded 99.7 mole %, and Koch-Light chloroform, which was additionally stabilized by addition of pentene-2.[19]

The ir spectra were recorded in the region of double-bond stretching frequencies, 1550–1750 cm^{-1}, in DMSO, and 1350–1750 cm^{-1} in D_2O as previously reported for other purine and pyrimidine derivatives,[5,17,20] and in the region of N–H and O–H stretches, 2900–3600 cm^{-1}, in $CDCl_3$, a solvent extensively used in studies on base pairing. Solubility limitations made it possible to record the neutral form in aqueous medium only for 4-thiouridine; for the remaining derivatives, the neutral forms were examined in DMSO and the uv spectra checked against those in H_2O to ensure that the same form existed in the aqueous and nonaqueous media. No difficulties were encountered with the solubilities of the anionic forms in aqueous medium.

Results and Discussion

We proceed first to an examination of the tautomeric forms of 4-methylthio-2-oxopyrimidine and its N_1-methyl derivative, both necessary as reference compounds.

4-Methylthio-2-oxopyrimidine may theoretically exist in one of three tautomeric forms (III–V) or a mixture



R = H, CH_3

of two or three of these. The pH-dependent uv absorption spectra are shown in Figure 1a and 1b. The equilibrium shown in Figure 1a clearly corresponds to protonation of a ring nitrogen and that in Figure 1b to ionization of the single potentially dissociable hydrogen. Figure 1c, in turn, exhibits the spectra of the neutral and protonated forms of 1-methyl-4-methylthio-2-oxopyrimidine, the structure of which is fixed as IV (R = CH_3).

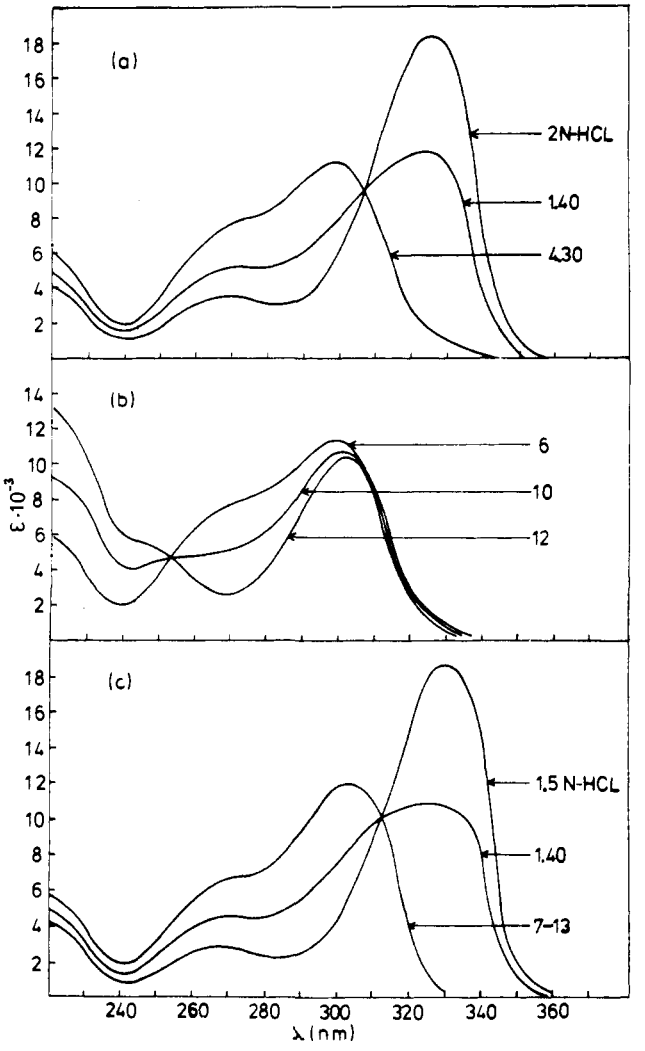


Figure 1. Ultraviolet absorption spectra at various pH values indicated beside each curve and showing: (a) the neutral (pH 4.3 curve) and ring N_3 protonated (2 *N* HCl curve) forms of 4-methylthiouracil, (b) the neutral (pH 6 curve) and anionic (pH 12 curve, dissociation of N_1 proton) forms of 4-methylthiouracil, (c) the neutral (pH 7–13 curve) and ring N_3 protonated (1.5 *N* HCl curve) forms of 1-methyl-4-methylthiouracil. For 3-methyl-4-methylthio-2-oxopyrimidine (spectra not shown), Maehr, *et al.*,[3] give the following data for λ_{max} (in nm) and ($\epsilon_{max} \times 10^{-3}$): in 0.1 *N* HCl 223 (16.6) 269 (4.75), 321 (17.3); at pH 7 and in 0.1 *N* KOH 226 (21.3), 320 (15.3). From this it is clear that these spectra are very different from those shown in (a) and (c).

From a comparison of Figure 1a with Figure 1c, it is clear that the structure of the neutral form of 4-methylthio-2-oxopyrimidine is IV. In agreement with this, the spectrally determined pK_a for protonation of 4-methylthio-2-oxopyrimidine is 1.55, as compared to 1.50 for 1-methyl-4-methylthio-2-oxopyrimidine (Table I), so that it is the N_3 nitrogen which undergoes protonation in IV. For the fixed reference compound 3-methyl-4-methylthio-2-oxopyrimidine, corresponding to III, the pK_a for protonation (in this case of the N_1 nitrogen) has been reported by Maehr, *et al.*,[3] as 4.07. The absorption spectra of the neutral and protonated forms of 3-methyl-4-methylthio-2-oxopyrimidine[3] also differ appreciably from those for 1-methyl-4-methylthio-2-oxopyrimidine and 4-methylthio-2-oxopyrimidine, as might be anticipated.

(15) A. Hampton and A. W. Nichol, *Biochemistry*, **5**, 2076 (1966).
(16) I. L. Doerr and J. J. Fox, *J. Org. Chem.*, **32**, 1462 (1967).
(17) Z. Kazimierczuk, A. Psoda, and D. Shugar, *Acta Biochim. Polon.*, 20, 83 (1973).
(18) S. Balslev, *Phys. Rev.*, **143**, 636 (1966).
(19) G. M. Nagel and S. Hanlon, *Biochemistry*, **11**, 816 (1972).
(20) H. T. Miles, *Proc. Nat. Acad. Sci. U. S.*, 47, 791 (1961); M. Tsuboi, *Appl. Spectrosc. Rev.*, 3, 45 (1969); A Psoda and D. Shugar, *Biochim. Biophys. Acta*, 247, 507 (1971).

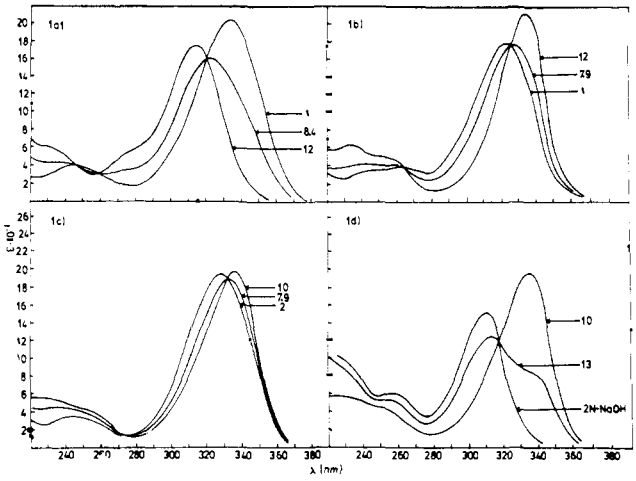


Figure 2. Ultraviolet absorption spectra at indicated pH values of the neutral and anionic forms of 4-thiouracil and its N-methylated derivatives: (a) 1-methyl-4-thiouracil, neutral (pH 1) and monoanionic (pH 12) forms, (b) 3-methyl-4-thiouracil, neutral (pH 1) and monanionic (pH 12) forms, (c) 4-thiouracil, neutral (pH 2) and monoanionic (pH 10) forms, (d) 4-thiouracil, monoanionic (pH 10) and dianionic (2 *N* HaOH) forms. See Table I for pK values obtained by spectral titration.

Table I. Spectrally Determined pK_a Values at about 25° for 4-Thiouracil Derivatives

Derivative	pK_1	pK_2
Uracil	9.5[a]	>13
4-Thiouracil	8.00	~12.8
5,6-Dihydro-4-thiouracil	9.8	
1-Methyl-5,6-dihydro-4-thiouracil	10.1	
1-Methyl-4-thiouracil	8.40	
3-Methyl-4-thiouracil	8.12[b]	
4-Methylthiouracil	1.55	9.9
1-Methyl-4-methylthiouracil	1.50	
3-Methyl-4-methylthiouracil	4.07[c]	
Uridine	9.25[d]	
4-Thiouridine	8.4[e]	

[a] From D. Shugar and J. J. Fox, *Biochim. Biophys. Acta*, **9**, 199 (1952). [b] Reference 3 gives 8.2. [c] From ref 3. [d] From J. J. Fox and D. Shugar, *Biochim. Biophys. Acta*, **9**, 369 (1952). [e] From B. C. Pal, *J. Org. Chem.*, **36**, 3026 (1971).

The infrared data are consistent with the assignment of 4-methylthio-2-oxopyrimidine as IV. In DMSO this compound, and the fixed 1-methyl-4-methylthio-2-oxopyrimidine, both exhibit $C_5{=}C_6$ band frequencies at 1616 cm^{-1} and carbonyl frequencies which differ only slightly as a result of N_1 methylation (Table II).

Uv Spectra of Neutral and Monoanionic Forms of 4-Thiouracils. 1-Methyl-4-thiouracil possesses only one potentially dissociable proton, and an examination of its uv spectrum as a function of pH (Figure 2a) shows, in fact, only one equilibrium with a pK_a value of 8.40. For purposes of simplicity, both in this and subsequent figures, we present only the spectra for the two extreme forms (in this case the neutral and monoanion) and one intermediate form at a pH in the neighborhood of the pK_a value.

For 3-methyl-4-thiouracil, which also has only one dissociable proton, the pH-dependent spectra (Figure 2b) again show only one equilibrium, the pK_a value for which was calculated as 8.12.

In the case of the parent 4-thiouracil, there are two ionizable hydrogens. In agreement with this, the pH dependence of the spectrum shows the presence of two equilibria, each with appropriate isosbestic points, exhibited in Figure 2c and Figure 2d. The former corresponds to monoanion formation with a pK_a of 8.00 and the latter to conversion of the monoanion to a dianion with a pK_a ~ 12.8.

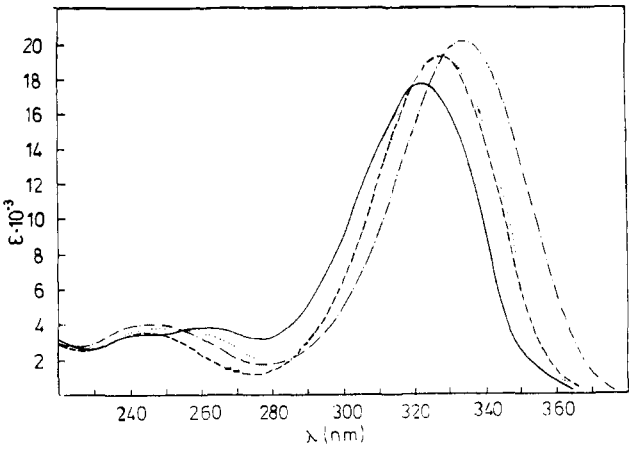


Figure 3. Ultraviolet absorption spectra of the neutral forms of (----) 4-thiouracil, (-·-·-) 1-methyl-4-thiouracil, (——) 3-methyl-4-thiouracil, and (·····) 1,3-dimethyl-4-thiouracil.

Neutral Forms of 4-Thiouracil Derivatives. Figure 3 exhibits the uv spectra for the neutral forms of 4-thiouracil and its monomethylated derivatives, taken from Figure 2a–c and for 1,3-dimethyl-4-thiouracil, the structure of which is fixed as the 2-keto-4-thione. Even a cursory examination demonstrates that all four compounds must be in the 2-keto-4-thione form. Note in particular that, relative to 4-thiouracil, N_1 methylation results in a bathochromic shift and *hyper*chromicity of the principal long-wavelength band, while N_3 methylation gives a hypsochromic shift with some hypochromicity of the long-wavelength band. The additivity of these two effects is testified to by the spectrum for 1,3-dimethyl-4-thiouracil, which is located midway between the two foregoing spectra and, consequently, virtually coincides with that for 4-thiouracil itself.

In accordance with the foregoing, 4-thiouracil and its N methylated derivatives all exhibit two characteristic strong infrared absorption bands in DMSO,[21] in the region embracing double-bond stretching frequencies (Figure 4 and Table II). The higher frequency band (1718 cm^{-1} in the case of 4-thiouracil) corresponds to the $C_2{=}O$, as in other purines and pyrimidines.[20] The lower frequency band (1616 cm^{-1} for 4-thiouracil) is due to the $C_5{=}C_6$ bond, confirmed by its disappearance in the spectrum of 1-methyl-5,6-dihydro-4-thiouracil, which exhibits only the $C_2{=}O$ band at 1707 cm^{-1} (Table II).

Particular attention is drawn to the very high integral intensity of the $C_5{=}C_6$ band, comparable to that for $C_2{=}O$. In uracil derivatives the intensity of the $C_5{=}C_6$ band is so low that it is frequently not detectable, probably due to the observed pronounced conjugation between $C_4{=}O$ and $C_5{=}C_6$.[8]

(21) The spectra of the neutral forms were run in DMSO since all the compounds, with the exception of 4-thiouridine, were insufficiently soluble in D_2O. Figures 4b and c exhibit the spectra of the neutral form of 4-thiouridine in D_2O and DMSO, respectively. The differences between these spectra are due to the solvent effect (see also Table II).

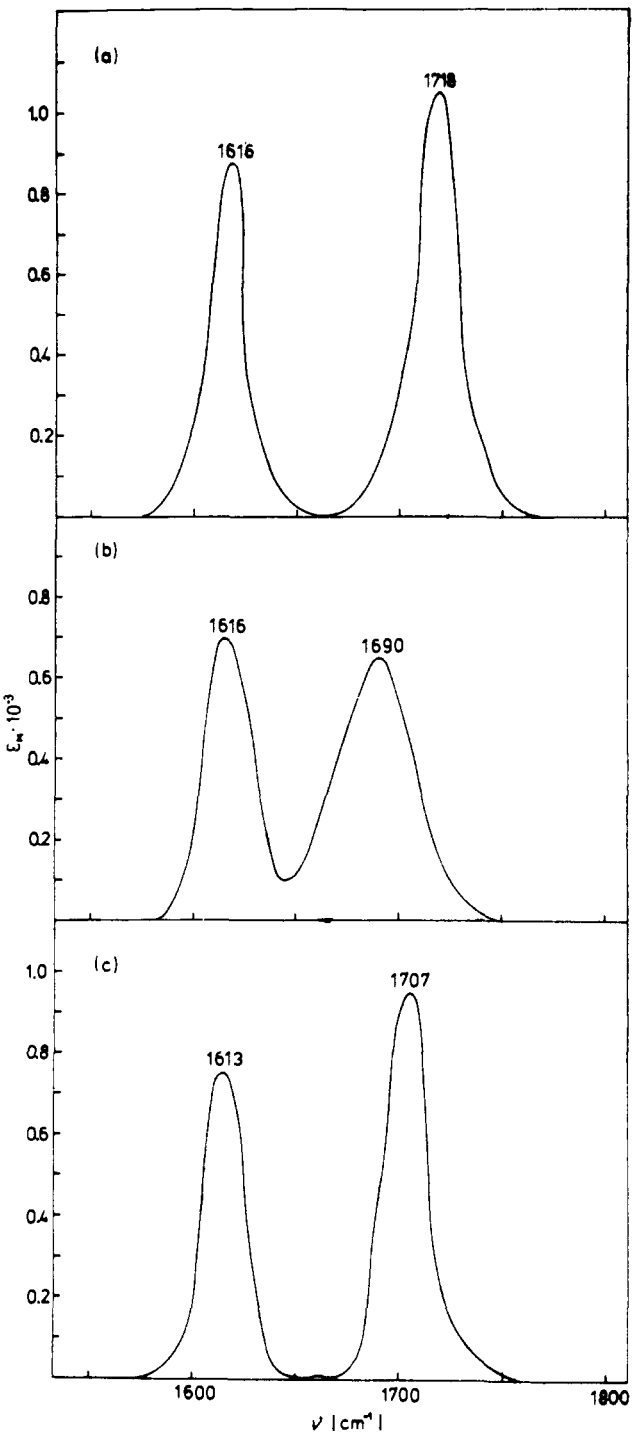


Figure 4. Infrared absorption spectra of (a) 4-thiouracil in DMSO, (b) 4-thiouridine in D_2O, and (c) 4-thiouridine in DMSO.

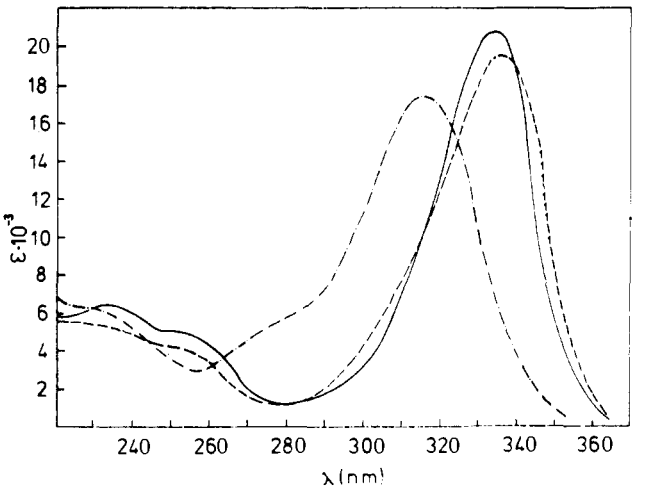


Figure 5. Uv absorption spectra (pH 10.5) of the monoanionic forms of (····) 4-thiouracil, (–·–·–) 1-methyl-4-thiouracil, and (——) 3-methyl-4-thiouracil. As pointed out in text, the near coincidence of the spectra for the monoanions of 4-thiouracil and 3-methyl-4-thiouracil is fortuitous and due to the (unknown) effect of N methylation.

If we compare the ir data of 4-thiouracil (Figure 4, Table II) and 4-methylthiouracil (Table II), it will be noted that the 1616-cm^{-1} band assigned to $C_5{=}C_6$ is the same for both, whereas the $C_2{=}O$ band in the latter is displaced 46 cm^{-1} to lower frequencies. For the corresponding 1-methyl derivatives the $C_5{=}C_6$ band of the *S*-methyl analog is displaced by 4 cm^{-1} and the C=O band by 43 cm^{-1}.

The foregoing observations further support the thione structure of the neutral forms of the 4-thiouracil derivatives, as well as 4-thiouridine, for the *S*-methyl derivatives are formally equivalent to the thiol forms. The low frequency of the $C_2{=}O$ band is characteristic for a pyrimidine ring in which this bond is conjugated to a 3,4-double bond, *e.g.*, cytosine,[20] 2-oxopyrimidine,[8] and 2-oxopurine.[22] The higher frequency of the $C_2{=}O$ band in 4-thiouracil is therefore consistent with its being in the thione form. Furthermore, if 4-thiouracil existed as an equilibrium mixture of the thione and thiol forms, one would anticipate the presence of two reasonably well-resolved bands in the C=O region, which is clearly not the case. Additional evidence for the 2-keto-4-thione form was forthcoming from an examination of the spectrum of 1-cyclohexyl-4-thiouracil in chloroform, in the region of the N–H and O–H stretches. A single N–H band was observed at 3385 cm^{-1}; no band corresponding to O–H in the 3600-cm^{-1} region was observable. The double-bond stretching frequencies were the same as in DMSO (Table II), allowing for the shifts normally encountered on transfer to a solvent with lower dielectric constant.

Neutral Form of 1-Methyl-4-thio-5,6-dihydrouracil. For the neutral form of 1-methyl-5,6-dihydro-4-thiouracil the 2-keto-4-thione structure is also indicated by the location of its carbonyl band at 1707 cm^{-1}. For if the structure were 2-keto-4-thiol, the conjugation of the $C_2{=}O$ with the $N_3{=}C_4$ bond would have resulted in an observed frequency for the former of about 1660–1670 cm^{-1}. It follows from the foregoing that 5,6-dihydro-4-thiouridine must also possess the 2-keto-4-thione structure.[23]

Monoanionic Forms. Figure 5 exhibits the uv spectra of the monoanionic forms of 4-thiouracil and its N_1 and N_3 methylated derivatives. The marked spectral differences between the latter two are consistent with dissociation of the N_3 proton in 1-methyl-4-thiouracil and the N_1 proton in 3-methyl-4-thiouracil. The

(22) L. M. Stempel, Ph.D. Thesis, Cornell University, 1968. We are indebted to Dr. J. J. Fox for making available the pertinent data from this thesis.

(23) CNDO/2 calculations of the electronic structure and total energies of thiouracils, carried out in this laboratory (M. Geller, A. Pohorille, and A. Jaworski, *Biochim. Biophys. Acta*, 331, 1 (1973)), are qualitatively in agreement with the 2-keto-4-thione structure as the most stable form for 4-thiouracil.

Table II. Frequencies, Integral Intensities, and Molar Extinction Coefficients of C_2=O and C_5=C_6 Bands of Neutral Forms of Various 4-Thiouracil Derivatives in Solvents as Indicated

Compound	Solvent	C_2=O (cm^{-1})	$10^{-4}A$ (M^{-1} cm^{-1})	ϵ_M (M^{-1} cm^{-1})	C_5=C_6 (cm^{-1})	$10^{4}A$ (M^{-1} cm^{-1})	ϵ_M (M^{-1} cm^{-1})
4-Thiouracil	DMSO	1718	9.6	1050	1616	5.6	920
1-Methyl-4-thiouracil	DMSO	1703	7.3	1210	1620	6.2	1080
1-Cyclohexyl-4-thiouracil	DMSO	1701	8.0	930	1611	4.7	930
	$CHCl_3$	1705	10.8	890	1617	6.3	830
3-Methyl-4-thiouracil	DMSO	1709	8.3	1100	1617	6.6	980
1,3-Dimethyl-4-thiouracil	DMSO	1690	6.5	1000	1624	6.0	980
	$CHCl_3$	1693	8.2	850	1630	4.9	1120
4-Thiouridine	DMSO	1707	6.5	1050	1613	5.2	900
	D_2O	1690	10.1	650	1616	6.2	690
1-Methyl-5,6-dihydro-4-thiouracil	DMSO	1707	6.9	900			
4-Methylthio-2-oxopyrimidine	DMSO	1672	8.5	1110	1616	4.3	810
1-Methyl-4-methylthio-2-oxopyrimidine	DMSO	1660	10.0	1390	1616	3.4	640
	$CHCl_3$	1655	12.5	1380	1622	4.7	650
O^2,2′-Anhydro-4-thiouridine	DMSO				1633	4.5	700

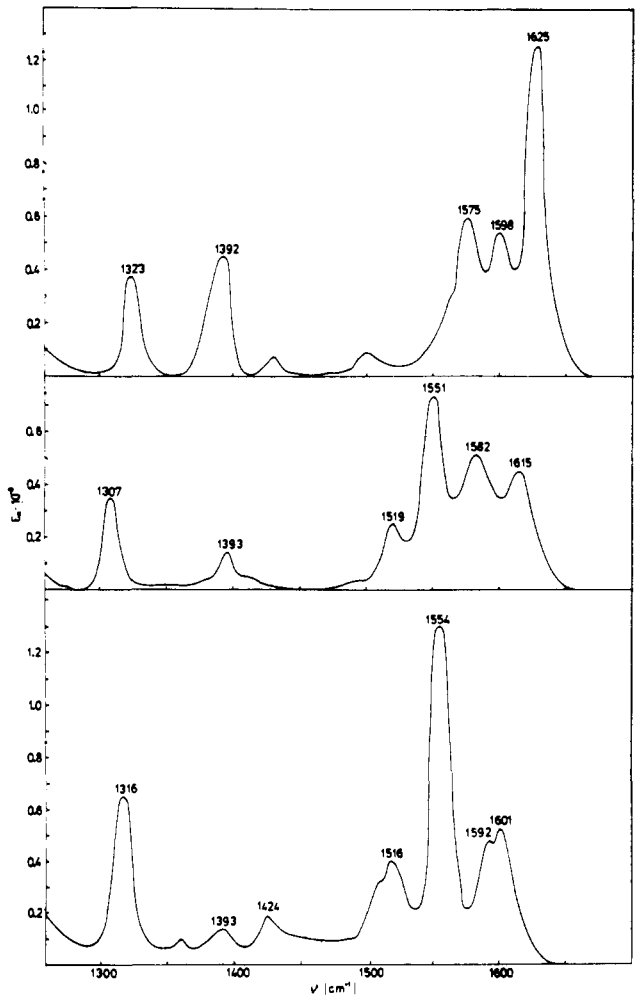


Figure 6. Infrared absorption spectra in D_2O of the monoanions of (a) 1-methyl-4-thiouracil, pD ~11 (the spectrum of 4-thiouridine under these conditions is very similar), (b) 4-thiouracil, pD ~10.5, (c) 3-methyl-4-thiouracil, pD ~11.

striking similarity between the monoanionic forms of 4-thiouracil and 3-methyl-4-thiouracil suggests, at first sight, that in the former it is also the N_3 hydrogen which ionizes to give the monoanion. This apparent resemblance is, however, purely coincidental in the light of the spectral shifts of N methylation on the neutral forms (Figure 3). A more logical interpretation is forthcoming from a comparison of the spectral changes accompanying ionization for 4-thiouracil and its two monomethylated derivatives (Figures 2a–c). Examination of Figure 2a and b shows that the very marked spectral changes accompanying monoanion formation in the two monomethylated derivatives are in diametrically opposing directions. This explains the relatively minor spectral modifications accompanying monoanion formation in 4-thiouracil itself (Figure 2c), which must consist of an *equilibrium*[24] mixture of the two types of monoanions, with resultant partial superposition of the two opposing spectral effects. From a visual examination of Figure 2a–c it may be deduced qualitatively that the proportions of the two monoanions in ionized 4-thiouracil are such that the form represented by ionized 3-methyl-4-thiouracil predominates. Information of a more quantitative nature is forthcoming from an examination of the first derivative spectrum of 4-thiouracil monoanion and the infrared spectra of the monoanions of 4-thiouracil and its N-methylated derivatives (see below).

It is now pertinent to inquire whether the structures of the monoanions of the foregoing derivatives involve localization of the negative charge on the O^2 or S^4 or charge delocalization as in the case of 2,4-diketopyrimidines (eq 1)[8] and as proposed for the corresponding monoanions of 2-thiouracil[25] and 2-thiobarbiturates.[17] We turn therefore to an examination of the infrared spectra of the monoanionic forms.

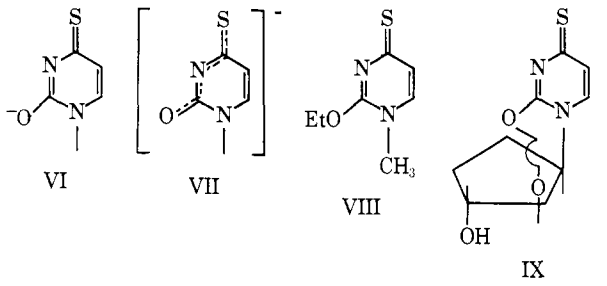
Ir Spectra and Structures of Monoanions. (a) 1-Methyl-4-thiouracil and 4-Thiouridine. The spectra of the monoanions of both these compounds (Figure 6a, Table II) in the region of double-bond stretching frequencies are very similar, as anticipated, and characterized by the absence of the carbonyl bands, and the presence of C_5=C_6 with frequencies 4 cm^{-1} higher than for the neutral forms. There are two additional bands, 1575 and 1598 cm^{-1} for 1-methyl-4-thiouracil and 1575 and 1600 cm^{-1} for 4-thiouridine, the assignments of which will be discussed below.

The disappearance in both monoanions of the C_2=O band, and maintenance of the C_5=C_6 band, is con-

(24) The existence of an *equilibrium* mixture of the two monoanions, both with the same pK values, follows from the well-defined isosbestic points at 275 and 332 nm in Figure 2c.

(25) D. Shugar and J. J. Fox, *Bull. Soc. Chim. Belg.*, 61, 293 (1952).

sistent with either of the structures VI or VII. Distinction between these two would be feasible with the aid of the reference compound 1-methyl-2-ethoxy-4-thiopyrimidine (VIII). This derivative is unknown, and recourse was then had to O^2,2′-anhydro-4-thiouridine (IX), which is formally identical with VIII.[26]



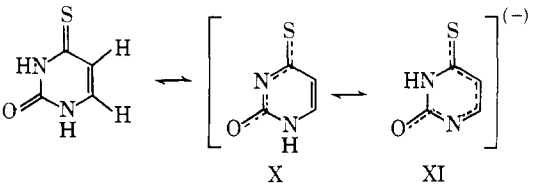
VI VII VIII IX

The spectrum of this derivative in DMSO (Table II) exhibits an intense band at 1633 cm^{-1}, corresponding to $C_5{=}C_6$, and a weaker band at 1524 cm^{-1}, probably C=N.[30] The appreciable spectral differences between O^2,2′-anhydro-4-thiouridine and the monoanion of 1-methyl-4-thiouracil consequently exclude VI as the structure of the latter and indicate that this is VII, with charge delocalization between O^2 and S^4.

(b) 3-Methyl-4-thiouracil. The ir spectrum of the monoanion of this derivative (Figure 6c) is characterized by the absence of the $C_2{=}O$ band and the appearance of new bands at 1516, 1554, and 1601 cm^{-1}, the latter of which is a superposition of two frequencies. The disappearance of the $C_5{=}C_6$ band is consistent only with the structure represented by I (eq 1), with charge delocalization between O^2 and S^4 *via* the C_5–C_6 bond.

(c) 4-Thiouracil. The monoanion of 4-thiouracil (Figure 6b) is devoid of the carbonyl frequency and exhibits new bands at 1519, 1551, and 1582 cm^{-1}. The band at 1615 cm^{-1} undoubtedly corresponds to the 1616-cm^{-1} band of the neutral form of 4-thiouracil in DMSO (Figure 4a) with the assignment $C_5{=}C_6$. Its integral intensity is, however, appreciably lower than that of the corresponding band in the monoanion of 1-methyl-4-thiouracil (Figure 6a). Furthermore the intense band at 1551 cm^{-1} corresponds to the 1554-cm^{-1} band of the monoanion of 3-methyl-4-thiouracil (Figure 6c). The infrared data therefore support the conclusions derived from the analysis of the uv spectra as to the existence of an equilibrium mixture of two monoanionic forms of thiouracil.

Furthermore the infrared data demonstrate that each of the two monoanionic forms involves extensive charge delocalization, as in the case of 2,4-diketopyrimidine,[8] so that the dissociation of the neutral form of 4-thiouracil may be represented as in eq 2.



X XI (2)

Tentative Infrared Band Assignments. The charge delocalization accompanying monoanion formation in 1-methyl-4-thiouracil occurs in the system $(O{=}C_2{-}N_3{=}C_4{-}S)^-$, as in the monoanion of 1-methyluracil;[8] however, the replacement of O^4 by S^4 necessarily leads to a reduction in symmetry of the system and a resultant different distribution of charge density.[31] For the monoanions of 1-methyluracil and 1-methylthymine, it proved possible to determine the band frequencies and band assignments of the system $(O{=}C_2{-}N_3{=}C_4{-}O)^-$ with the aid of appropriate model systems in which the 5,6-bonds were saturated.[8,32] Since no such model systems are available for 4-thiouracils, we have attempted to make some band assignments by comparison with those reported for the monoanion of 1-methyluracil.[8]

The shortest wavelength band of the 1-methyl-4-thiouracil monoanion, 1625 cm^{-1} (ϵ_M ~1360), corresponds to $\nu(C_5{=}C_6)$. For 1-methyluracil monoanion this band is located at 1638 cm^{-1} (ϵ_M ~900). The difference in intensities is due to the low intensity of this band in the neutral forms of 2,4-diketopyrimidines[8] and its relatively large intensity in the neutral forms of 2-keto-4-thiopyrimidines.

The doublet at 1598 cm^{-1} (ϵ_M ~540) and 1575 cm^{-1} (ϵ_M ~600) in 1-methyl-4-thiouracil monoanion finds its counterpart in the doublet at 1598 cm^{-1} (ϵ_M ~500) and 1580 cm^{-1} (ϵ_M ~600) exhibited by 1-methyluracil monoanion and assigned to $\nu^{A_1}(C{=}N)$. No assignment was made for $\nu^{B_1}(C{=}N)$ in the 1-methyluracil anion, but this band was located at 1315 cm^{-1} (ϵ_M ~330) in the monoanion of the uracil photodimer. A similar band, 1325 cm^{-1} (ϵ_M ~380), is exhibited by 1-methyl-4-thiouracil monoanion.

In the spectrum of the 1-methyluracil monoanion, the two most intense bands at 1499 and 1450 cm^{-1} were identified as $\nu^{B_1}(C{=}O)$ and $\nu^{A_1}(C{=}O)$.[8] The 1392-cm^{-1} band in the spectrum of the 1-methyl-4-thiouracil monoanion possibly corresponds to $(C_2{=}O)^-$. However the accompanying $(C_4{=}S)^-$ frequency, like those

(26) The validity of the use of this reference compound is supported by the following. The uv spectrum of the neutral form of the known N_1-riboside of 2-ethoxy-4-pyrimidone exhibits maxima at 225 and 250 nm and a minimum at 236 nm,[27] while that of O^2,2′-anhydrouridine has maxima at 223 and 250 nm and a minimum at 236 nm.[28,29] The close similarity between these spectra is due to the relatively low strain in the heterocyclic ring of the O^2,2′-derivative. For the corresponding O^2,5′-derivatives the strain in the heterocyclic ring is more pronounced with concomitant appreciable differences in the uv spectra.[27]

(27) J. Pitha, *J. Org. Chem.*, **35**, 903 (1970).

(28) D. H. Shannahoff and R. A. Sanchez, *J. Org. Chem.*, **38**, 593 (1973).

(29) J. J. Fox and I. Wempen, *Advan. Carbohyd. Chem.*, **14**, 282 (1959).

(30) The 4-keto derivative O^2,2′-anhydrouridine exhibits, in addition to the carbonyl band at 1649 cm^{-1}, a weak band visible as a shoulder at 1633 cm^{-1}, assigned to $C_5{=}C_6$. Hence, in this system as well, replacement of the exocyclic $=O^4$ by $=S^4$ leads to a marked increase in the $C_5{=}C_6$ band intensity.

(31) The charge distribution densities for the two monoanions of thymine have been derived by all-electron all-integral LCAO–SCF calculations (L. C. Snyder, R. G. Shulman, and D. B. Neumann, *J. Chem. Phys.*, **53**, 256 (1970)). To our knowledge, no such calculations have yet been made for the analogous 4-thiothymine or 4-thiouracil.

(32) Ideal model systems are obviously the corresponding 5,6-dihydro derivatives. However, since the monoanions of 5,6-dihydro-2,4-diketopyrimidines are normally alkali labile, the 5,6;5′,6′-cyclobutane photodimers of uracil and thymine were employed in the foregoing study.[8] The photodimers, which contain 5,6-saturated bonds, are relatively stable in alkali. We have since further confirmed the validity of the use of the photodimers as a result of the finding that the monoanion of 1,6-dimethyl-5,6-dihydrouracil is remarkably stable in alkali and exhibits infrared band frequencies analogous to those of the monoanions of 1-methyluracil and 1-methylthymine (K. L. Wierzchowski, E. Litońska, and D. Shugar, unpublished). Unfortunately 4-thiouracils to not form cyclobutane photodimers; furthermore the monoanions of their 5,6-dihydro derivatives proved too alkali labile to make possible measurements of infrared band frequencies.

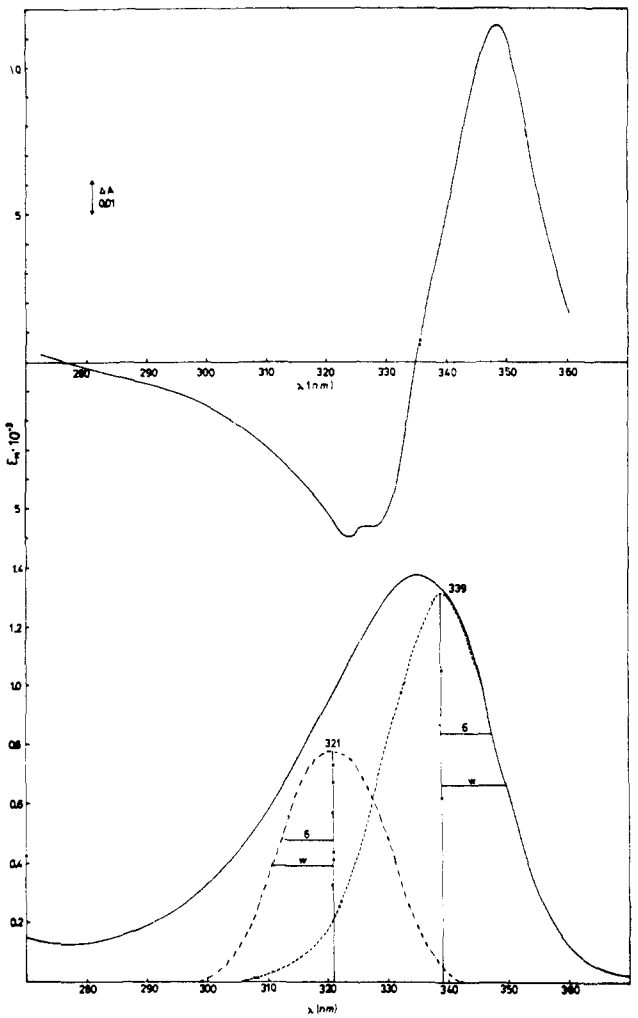


Figure 7. (a) First derivative uv spectrum of 4-thiouracil monoanion, showing (in region between 320 and 330 nm) the presence of two monoanionic species; (b) resolution of the individual bands for the two monoanions according to the procedure of Shiga, *et al.*,[34] with L_1 = 348 nm, L_2 = 353 nm, ΔL = 5 nm, σ = 9 nm, W = 11 nm, and c = 339 nm (see ref 34 for further details of procedure).

associated with C=S and C—SH bonds,[33] would probably not be visible in the infrared under our experimental conditions. More complete and definitive band assignments for the compounds embraced in this study, as well as for thiopyrimidines and thiopurines in general, must await the application of Raman spectroscopy, now under way in this laboratory.

Tautomeric Proportions of 4-Thiouracil Monoanions. For the mixtures of monoanions of various 2,4-diketopyrimidines, the uv spectra of the two monoanionic species are sufficiently widely separated so that, with suitable corrections, the proportions of the two monoanionic species may be calculated from the extinctions of their long-wavelength bands,[6-8] the results being in reasonably good agreement with those obtained from the infrared spectra, where characteristic bands for each species are also well resolved.[8]

It is clear from Figure 5 that the uv spectrum of 4-thiouracil monoanion is of little value for even a rough approximation of the proportions of the two types of monoanions. Recourse was therefore had to first derivative spectroscopy, and the results were compared with those obtained from infrared spectroscopy.

(a) The first derivative spectrum of the 335-nm band of 4-thiouracil monoanion, shown in Figure 7a clearly demonstrates the existence of two bands corresponding to the two monoanionic species X and XI (eq 2). With the aid of this spectrum, the 335-nm absorption band of the 4-thiouracil monoanion was resolved by the procedure of Shiga, *et al.*,[34] with results shown in Figure 7b. From this figure, with due allowance for the asymmetry of the 335-nm band at its short wavelength extremity, the ratio of the two monoanions was estimated as X:XI ~ 1:2.7.[35]

(b) A reasonable estimate of the proportion of tautomer X in 4-thiouracil monoanion may be obtained from the ratio of the integral intensities of the (C_5=C_6) bands in the monoanions of 4-thiouracil and 1-methyl-4-thiouracil (Figure 6b and a). The ratio of the two monoanions obtained in this way is X:XI ~ 1:3, a result in reasonable agreement with that obtained from the uv derivative spectrum, above.

Solvent-Induced Shift of Tautomeric Equilibrium. For the monoanionic forms of uracil, thymine, and 5-fluorouracil, each of which exhibits an equilibrium mixture of monoanions I and II, a decrease in the dielectric constant of the solvent medium (with the use of aqueous dioxane) led to an increase in the proportion of the less polar monoanion II.[8] Similar shifts were observed for the monoanion equilibrium of 4-thiouracil in aqueous dioxane, but the results obtained were only qualitative because of solvent effects which were reflected in modifications of the spectra of the monomethylated derivatives.

Recourse was then had to a divalent cation, Ca^{2+}. It was noted some time ago by Clauwaert and Stockx[36] that addition of $CaCl_2$ to uracil monoanion led to pronounced changes in the uv absorption spectrum. This was ascribed to binding of Ca^{2+} by the uracil monoanion, but no such effect was observed with the monoanions of 1-methyluracil or 3-methyluracil. These authors were not aware at the time that uracil forms a mixture of two monoanions. A reexamination of their spectral data shows clearly that the equilibrium of the two monoanions is shifted in favor of monoanion I with an increase in Ca^{2+} concentration.[37]

Direct examination showed that elevated concentrations of $CaCl_2$ did not affect the uv spectra of either the neutral or monoanionic forms of 1-methyluracil and 3-methyluracil. By contrast the absorption spectrum of 4-thiouracil monoanion(s) is appreciably modified in the presence of increasing concentrations of $CaCl_2$ in a manner consistent with a shift in the tautomeric ratio of the two forms (Figure 8) in favor of monoanion X at the expense of XI. Note, in particular, the

(33) B. Schrader, *Angew. Chem., Int. Ed. Engl.*, 12, 884 (1973).

(34) T. Shiga, K. Shiga, and M. Kuroda, *Anal. Biochem.*, 44, 291 (1971).

(35) Derivative spectroscopy of the monoanions of 1-methyl-4-thiouracil and 3-methyl-4-thiouracil showed no splitting of the long-wavelength bands, in agreement with the existence of one species for each of these. Application of our system to thymine, for purposes of control, gave the expected splitting for a mixture of two monoanionic species in the ratio of 1:1.[6-8]

(36) J. Clauwaert and J. Stockx, *Z. Naturforsch. B*, 23, 30 (1968).

(37) This shift in the equilibrium concentration of the two monoanionic forms by Ca^{2+} has since been extended and confirmed for thymine, 6-methyluracil, and 5-fluorouracil (J. Clauwaert and D. Shugar, unpublished observations). See also ref 38 for influence of other cations and anions.

(38) R. Shapiro and S. Kang, *Biochim. Biophys. Acta*, 232, 1 (1971).

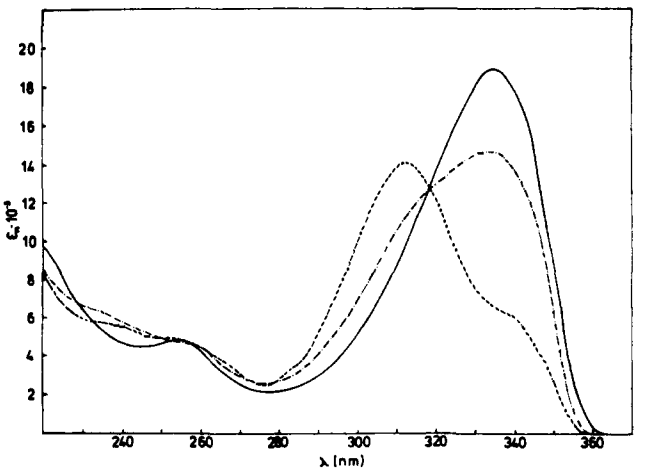


Figure 8. Ultraviolet absorption spectrum of 4-thiouracil monoanion(s) in D_2O and the Ca^{2+} induced shift in the ratio of the two monoanionic forms: (——) in ammonium buffer, pD ~10.5; (-·-·-) in the presence of 1 *M* $CaCl_2$, pD ~10.5; (- - - -) in the presence of 5 *M* $CaCl_2$, pD ~10.5.

isosbestic point at 318.5 nm in Figure 8, testifying to a shift in equilibrium between the two forms.[39]

In the infrared spectrum of 4-thiouracil monoanion, addition of 1 *M* $CaCl_2$ led to an increase in the integral intensity of the 1615-cm^{-1} band (characteristic for the 1-methyl-4-thiouracil monoanion) from 455 to 505 M^{-1} cm^{-1} and a decrease in intensity of the 1554-cm^{-1} band (characteristic of 3-methyl-4-thiouracil monoanion) from 755 to 620 M^{-1} cm^{-1} (Figure 9).[40] An increase in $CaCl_2$ concentration to 5 *M* led to a further increase in intensity of the 1615-cm^{-1} band to 580 M^{-1} cm^{-1} and to an additional decrease in that of the 1554-cm^{-1} band to 590 M^{-1} cm^{-1} (Figure 9).

It remains to clarify the source of the Ca^{2+}-induced shift in the equilibrium ratio of the two tautomeric monoanions. The decrease in dielectric constant of aqueous solutions of neutral salts is well documented,[41,42] but this effect alone cannot account for

(39) We have reexamined, in the light of these findings, our previous results for the monoanions of 2-thiouracil and its monomethylated derivatives (see ref 25) and, as might have been anticipated, find that 2-thiouracil also consists of an equilibrium mixture of two monoanionic forms (A. Psoda and D. Shugar, in preparation), a fact of some interest in relation to the pharmacological activity of 2-thiouracil.

(40) These modifications in integral intensities are not far outside the limits of experimental error of about 10%. However, the fact that the intensity of one band increases *simultaneously* with a decrease in the other and that both change in the direction to be anticipated from the changes in uv spectra (Figure 8) removes any doubts as to the validity of these observations, which were also found to be fully reproducible in several experiments.

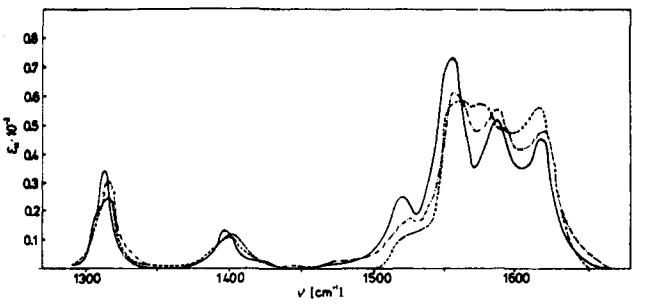


Figure 9. Infrared absorption spectrum of 4-thiouracil monoanions in D_2O: (——) pD ~10.5; (-·-·-) in the presence of 1 *M* $CaCl_2$, pD ~10.5; (- - - -) in the presence of 5 *M* $CaCl_2$, pD ~10.5.

the shift in equilibrium between the two monoanions, since in the case of thymine the shift is in favor of type II monoanion with aqueous dioxane and in favor of type I monoanion with $CaCl_2$ solutions. Additional factors must be involved, such as solute-induced changes in the structure of the solvent.[42-44]

4-Thiouracil Dianion. In the uv absorption spectrum of the dianion of 4-thiouracil, the long-wavelength maximum is appreciably blue-shifted with respect to the monoanion (Figure 2d), the resulting spectrum being similar to that of the monoanion of 4-methylthiouracil (Figure 1b), as might be anticipated from the fact that both of these arise from fully aromatic ring systems. In agreement with this, the infrared spectra in D_2O (not shown) are also very similar.

Acknowledgments. We are indebted to Professor M. Grynberg for supervision of the construction of the derivative spectrophotometer, to Elizabeth Olobry for the derivative spectra and their analyses, to J. Giziewicz for synthesis of the O^2,2′-anhydro nucleosides, and to Krystyna Macalska for assistance in recording some of the infrared spectra. This investigation was supported by the Polish Academy of Sciences (Project 09.3.1.), the Wellcome Trust, and the Agricultural Research Service, U. S. Department of Agriculture.

(41) R. A. Robinson and R. H. Stokes, "Electrolyte Solutions," Butterworth, London, 1959, pp 18–19.

(42) N. E. Hill, W. E. Vaughan, A. H. Price, and M. Davies, "Dielectric Properties and Molecular Behaviour," Van Nostrand–Reinhold, London, 1968, p 348 ff, and references therein.

(43) In an independent study on the tautomerism of the neutral forms of pyrimidines such as isocytosine and 2-ethoxy-4-pyrimidone, we have found similar shifts in the proportions of the two tautomers with aqueous dioxane and $CaCl_2$, to be reported elsewhere.

(44) Since completion of this study, we have found that the ratio of the two monoanionic forms of 4-thiouracil may be readily estimated by measurements of the chemical shifts of H_5, which are different for the two monoanionic species, as might have been anticipated. Further studies on this are in progress.