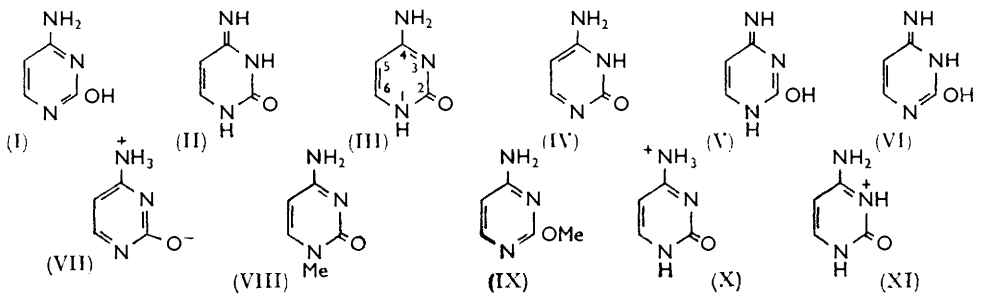
559. *Tautomeric Azines. Part III.*[1] *The Structure of Cytosine * and its Mono-cation.*†

By A. R. KATRITZKY and A. J. WARING.

Structures (III) and (XI) respectively are confirmed to predominate for cytosine and its cation in both aqueous and dimethyl sulphoxide solutions. The equilibrium constant for tautomerisation to the alternative amino-oxo-form (IV) is $10^{2\cdot9}$ by pK_a measurements. The 1,N^4-dimethylcytosine cation shows nuclear magnetic resonance spectra for the individual *cis* and *trans* single-bond isomers (XXVII) and (XXVIII).

CYTOSINE is important as a constituent of all nucleic acids. It could exist as seven ‡ tautomers: one aminohydroxy- (I), one imino-oxo- (II), two amino-oxo- (III, IV), two hydroxy-imino- (V, VI), and one betaine (VII) form. The actual structure of cytosine has been much discussed. Early infrared work by Blout and Fields [2] appeared to indicate that it existed in the hydroxy-form (I), and this received apparent confirmation by Stimpson and O'Donnell [3] on the basis of infrared and ultraviolet spectra. However, later ultraviolet (and infrared [5]) data have pointed towards the predominance of form (III). Shugar and Fox [6] showed that its spectra resembled those of structure (VIII) and not (IX), thus ruling out the possibility that cytosine existed in a hydroxy-form, and Kenner, Reese, and Todd [7] similarly established that imino-forms were unimportant. Early workers [8] considered that the cytosine cation existed as (X). However, protonation to give structure (XI) is much more likely in view of the behaviour of analogous heterocyclic compounds, and Dekker [9] has presented ultraviolet evidence for this view.



Recently, Kokko, Goldstein, and Mandell [10] suggested structures (VII) and (X) for cytosine and its cation on the basis of proton resonance spectroscopy carried out in deuterated dimethyl sulphoxide. As these conclusions seemed likely to be accepted,[11]

* Cytosine derivatives are numbered throughout this paper as in (III).

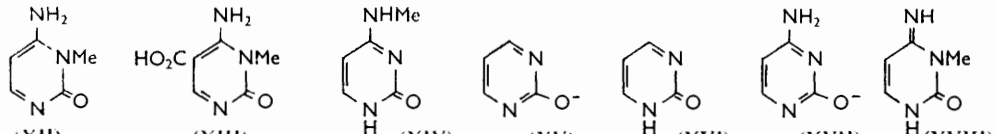
† Some of these conclusions have been given in a preliminary communication: Katritzky and Waring, *Chem. and Ind.*, 1962, 695.

‡ This excludes forms in which the aromaticity of the ring is totally lost, *i.e.*, in which one of the ring-atoms is sp^3 hybridised. Tautomerism to such forms has been suggested (*e.g.*, ref. 4), but it is very unlikely.

[1] Part II, Katritzky and Waring, *J.*, 1962, 1544.
[2] Blout and Fields, *J. Amer. Chem. Soc.*, 1950, **72**, 479.
[3] Stimson and O'Donnell, *J. Amer. Chem. Soc.*, 1952, **74**, 1805.
[4] Wang, *Nature*, 1959, **184**, 184.
[5] Angell, *J.*, 1961, 504.
[6] Shugar and Fox, *Biochim. Biophys. Acta*, 1952, **9**, 199.
[7] Kenner, Reese, and Todd, *J.*, 1955, 855.
[8] Sinsheimer, Nutter, and Hopkins, *Biochim. Biophys. Acta*, 1955, **18**, 13.
[9] Dekker, *Ann. Rev. Biochem.*, 1960, **29**, 453.
[10] Kokko, Goldstein, and Mandell, *J. Amer. Chem. Soc.*, 1961, **83**, 2909.
[11] Personal communications.

and as we felt that their results could be interpreted on the basis of formulæ (III) and (XI), we now offer additional evidence for the predominance of these structures.

Preparations.—In 1955, Whitehead and Traverso [12] described the preparation of 3-methylcytosine (XII), m. p. 260—265°, by the decarboxylation of the 5-carboxylic acid (XIII). However, this material was shown by Brown [13] to be 4-methylamino-2-pyrimidone (XIV) which must have resulted by rearrangement during the pyrolysis. In our hands pyrolysis at 10^{-2} mm. gave a sublimate of 3-methylcytosine, m. p. 213°, with ultraviolet spectra and pK values in agreement with material prepared by another route.[14] Through the kindness of Dr. P. D. Lawley a direct comparison was made which confirmed the identity. 2-Pyrimidone, prepared by the literature methods,[15,16] was obtained in a new polymorphic modification.



Ultraviolet Spectra.—The spectrum of cytosine resembles that of 1-methylcytosine (VIII), but differs from that of 3-methylcytosine (XII), and of the anion (XV) of 2-pyrimidone (Table 1, and Figs. 1 and 2). In conjunction with the previous ultraviolet evidence, these results confirm that cytosine exists as (III) and in particular show that (IV) and (VII) are unimportant in the tautomeric equilibrium, in both aqueous and dimethyl

TABLE 1.

Ultraviolet spectra and pK values at 23° ± 1°.

	Neutral			Cation			Anion			
	Solvent	$\lambda_{max.}$ (mμ)	$\varepsilon_{max.}$ × 10^{-3}	Solvent	$\lambda_{max.}$ (mμ)	$\varepsilon_{max.}$ × 10^{-3}	Solvent	$\lambda_{max.}$ (mμ)	$\varepsilon_{max.}$ × 10^{-3}	pK_a *
Cytosine	pH 7	267	6·13 †	0·1N-HCl	276	10·0 †	—	282	7·86 †	4·45 †
	Me_2SO	270	4·73	Me_2SO	282	10·9	—	—	—	4·60 [a]
1-Methylcytosine	pH 7	273	8·15	0·1N-HCl	282	12·2	—	—	—	4·55 †
	Me_2SO	278	6·8	—	—	—	—	—	—	4·57 [b]
3-Methylcytosine	0·01N-NaOH	293	12·5	pH 4	273	9·6	—	—	—	7·49 ± 0·04 [c]
2-Pyrimidone	pH 7	299	4·73	N-H_2SO_4	309	5·90	0·01N-NaOH	291	4·80	2·24 ± 0·04 [d]
	Me_2SO	311	3·98	—	—	—	Me_2SO	312	4·50	—

* Determined titrimetrically unless otherwise stated. † From ref. 6. pK_a determined spectrophotometrically.

[a] Levene, Bass, and Simms, *J. Biol. Chem.*, 1926, **70**, 229. [b] From ref. 7. [c] Titrated in 0·017M-solution at 20°. [d] From ref. 15.

sulphoxide solution, because the ultraviolet spectra of (IV) and (VII) would resemble those of (XII) and (XV), respectively. Further, the spectra of the cations of cytosine and its 1- and 3-methyl derivatives are similar, and different from the spectrum of 2-pyrimidone (XVI) (Table 1, Figs. 3 and 4), indicating that the first three cations all have structures of type XI and not of type X.

Basicity Measurements (Table 1).—Since cytosine and 1- and 3-methylcytosine are protonated to give cations of similar structure, the equilibrium constant for (III) ⇌ (IV) may be calculated as *ca.* $10^{2\cdot9}$, in favour of structure (III). This is comparable to the

[12] Whitehead and Traverso, *J. Amer. Chem. Soc.*, 1955, **77**, 5867.
[13] Brown, *J. Appl. Chem.*, 1955, **5**, 358.
[14] Lawley, personal communication; Brookes and Lawley, *J.*, 1962, 1348.
[15] Brown, *Nature*, 1950, **165**, 1010.
[16] Hunt, McOmie, and Sayer, *J.*, 1959, 525.

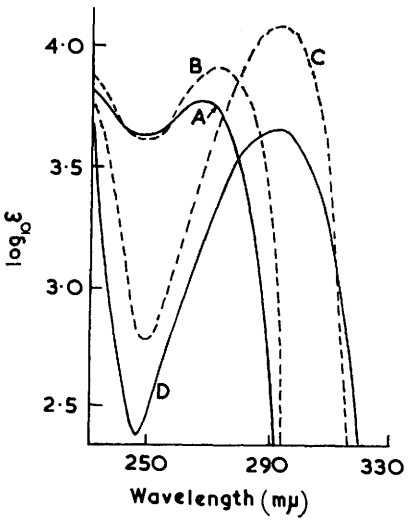


FIG. 1. In aqueous solution. Neutral molecules of cytosine (A), 1-methylcytosine (B), and 3-methylcytosine (C); anion of 2-pyrimidone (D).

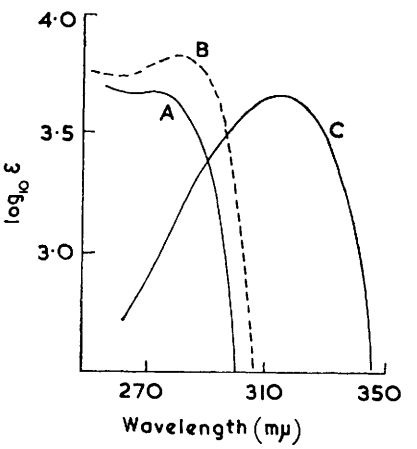


FIG. 2. In dimethyl sulphoxide. Neutral molecules of cytosine (A) and 1-methylcytosine (B); anion of 2-pyrimidone (C).

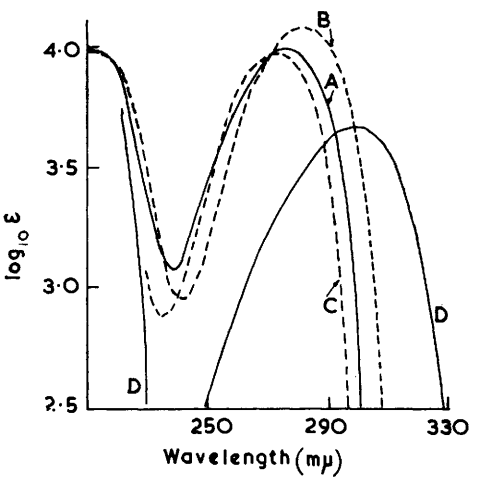


FIG. 3. In aqueous solution. Cations of cytosine (A), 1-methylcytosine (B), and 3-methylcytosine (C); neutral 2-pyrimidone (D).

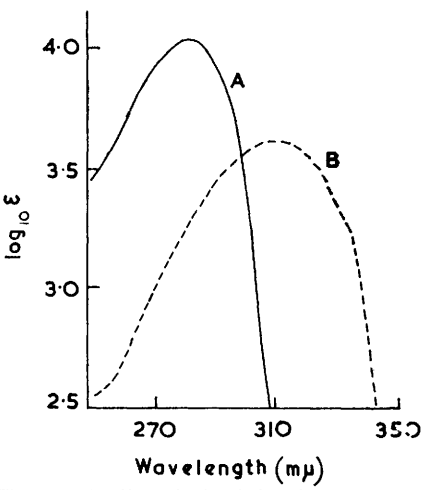


FIG. 4. In dimethyl sulphoxide. Cation of cytosine (A); neutral 2-pyrimidone (B).

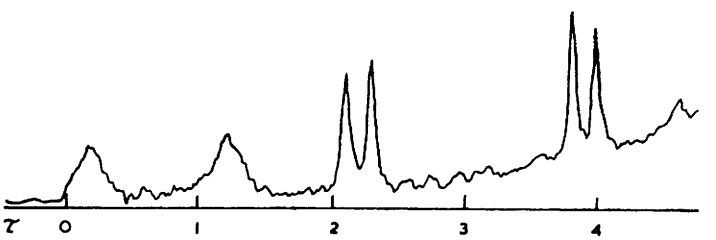


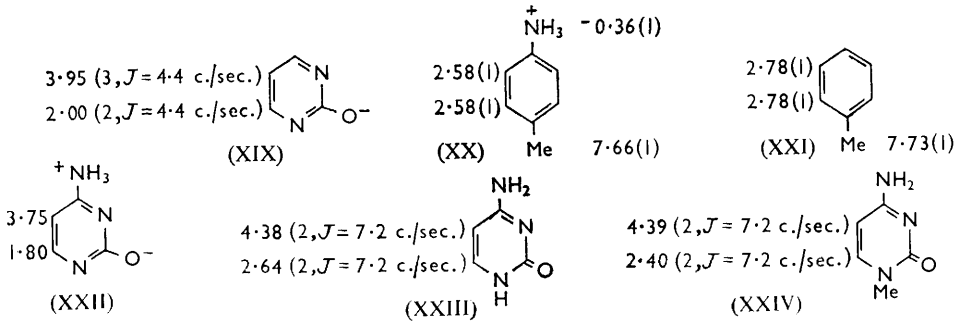
FIG. 5. Nuclear magnetic resonance spectrum of cytosine hydrochloride in dimethyl sulphoxide.

ratio of the K_A values for 2- and 4-aminopyridine; the relevance of this comparison is evident if the protonation of the common anion (XVII) is considered.

3-Methylcytosine (XII ⇌ XVIII) is shown to have a tautomeric constant of *ca.* 60, in favour of the amino-form (XII), by comparison of its pK value (7·49 ± 0·04) with that of 1,3-dimethylcytosine (pK = 9·29).[7]

Proton Resonance Spectra.—We measured the chemical shifts shown (τ values in p.p.m. relative to tetramethylsilane = 10; multiplicity and coupling constant shown in parentheses) in (XIX), (XX), and (XXI) in dimethyl sulphoxide solution, and thus calculated that species (XXII) should have the chemical shifts indicated. The chemical shifts determined for cytosine (XXIII) agreed with those for 1-methylcytosine (XXIV) and not with those calculated for species (XXII). In deuterium oxide the shifts for (XXIII) [τ 2·51, 4·06] and (XXIV) [τ 2·49, 4·11] were little changed. The chemical shift for cytosine did not vary with concentration.

Proton Magnetic Resonance of Cytosine Cations (Table 2).—In dimethyl sulphoxide, cytosine hydrochloride shows broad peaks at τ 0·15 and 1·20. These are each of relative area 1, and are assigned (not necessarily respectively) to the protons (a) and (b) in (XXV); R = H, which are non-equivalent because of restricted rotation about the NH_2–C bond



(cf. XXVI). This assignment is supported by the occurrence of these bands in the 1-methylcytosine cation (XXV; R = Me). Further support is afforded by the 1,N^4-dimethylcytosine cation: this shows *two* AB patterns for the 5- and 6-position protons, which we assign to the *cis–trans* isomers (XXVII) and (XXVIII) present in unequal proportions (the 5- and 6-protons form a single AB system for all the other compounds

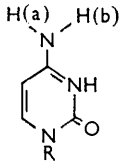
TABLE 2.

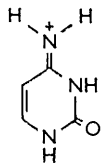
Proton magnetic resonance spectra of cytosine cations.

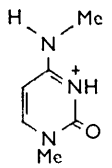
Compound [a]	Solvent	5-H (τ)	6-H (τ)	J_{56} (c./sec.)	N_1–Me (τ)	N_4–Me (τ)	NH_2 or NHMe (τ)
Cytosine	Me_2SO	3·90	2·20	8·0	—	—	0·15, 1·20
	D_2O	3·85	2·26	8·0	—	—	—
1-Methylcytosine	Me_2SO	3·80	1·90	8·0	*ca.* 6·6	—	0·05, 1·15
	D_2O	3·80	2·10	7·5	6·54	—	—
	SO_2	3·61	2·13	8·0	6·44	—	*ca.* 0·8 (v.b.)
1,N^4-Dimethylcytosine	Me_2SO	3·63, 3·67	2·00	8·0	6·62	—	—
	H_2SO_4	—	1·97, 2·21	8·0, 8·0	6·54	6·91 (5 c./sec.)	—
	SO_2	3·78, 3·54	2·07, 2·29	8·0, 8·0	6·46	6·82 (5 c./sec.), 6·75 (5·5 c./sec.)	—
	D_2SO_4	3·70, 3·80	2·00, 2·24	8·0, 8·0	6·56, 6·54	6·90	—
1,N^4,N^4-Trimethylcytosine	Me_2SO	3·52	1·72	8·0	*ca.* 6·5	*ca.* 6·5	—
	SO_2	3·61	2·10	8·0	6·44	6·50, 6·55	—
	D_2SO_4	3·69	2·16	8·0	6·56	6·68	—

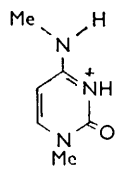
[a] Hydrochlorides except for solutions in deuterium sulphate.

investigated). In liquid sulphur dioxide, there is a difference in chemical shift between the 4-methyl group in the isomers (XXVII) and (XXVIII), and each methyl group is


(XXV)


(XXVI)


(XXVII)


(XXVIII)

spin-spin coupled to the adjacent NH; however, in deuterium sulphate the chemical shift between these 4-methyl groups must be very small. 1,N^4,N^4-Trimethylcytosine cation similarly shows the chemical-shift difference between the two N^4-methyl groups in sulphur dioxide but not in deuterium sulphate. In 4N-sulphuric acid, 1,N^4-dimethylcytosine shows *two* peaks for the N^4-methyl group, due to spin-spin coupling with the NH; the doublet condenses into a singlet in concentrated sulphuric acid, probably because of double protonation (similar results have been found for methylamidines [17]).

As a further check we have measured the spectrum of 1,N^4-dimethylcytosine hydrochloride and 1,N^4,N^4-trimethylcytosine hydrochloride at 60 Mc./sec.* The results fully confirm the assignments suggested: chemical shifts were within *ca.* 0·05 p.p.m. and coupling constants within *ca.* 1 c./sec. of the values at Mc./sec.

Conclusion.—The above evidence shows clearly that cytosine exists predominantly as 4-amino-2(1*H*)-pyrimidone, and that it undergoes protonation at the 3-position. Proton magnetic resonance spectroscopy has rapidly become one of the most powerful tools available for structure determination: we have elsewhere called attention to limitations inherent in its application.[18]

Experimental

1-Methylcytosine (m. p. 299—300, lit.,[7] m. p. 300°) was prepared by the literature method.

5-Ethoxycarbonyl-3-methylcytosine.—*N*-Methyl-*N'*-(2-cyano-2-ethoxycarbonylvinyl)urea [12] (15 g.) was refluxed for 2 hr. with ethanolic sodium ethoxide [from sodium (1·75 g.) and ethanol (40 c.c.)] and the whole evaporated at 100°/15 mm. Acetic acid was added to a solution of the residue in water until the pH was 5, giving the ester (11·0 g., 74%), m. p. 232—234° (decomp.) (from ethanol) (lit.,[12] m. p. 230°). This represents an improvement on the synthesis described in ref. 12.

3-Methylcytosine-5-carboxylic Acid.†—5-Ethoxycarbonyl-3-methylcytosine (2·0 g.) was heated at 100° for 3 hr. with N-sodium hydroxide (15 c.c.). Acidification of the cooled solution with acetic acid gave 3-methylcytosine-5-carboxylic acid (1·2 g., 70%). After recrystallisation from water (charcoal) the m. p. was 248—249° (decomp.) [lit., 251° (decomp.); [12] 240° (decomp.) [13]].

3-Methylcytosine.—3-Methylcytosine-5-carboxylic acid (0·172 g.), m. p. 248—249° (decomp.), was heated at 252—255°/0·01 mm. The yellow sublimate (40·4 mg., 32%) was recrystallised from aqueous acetone (charcoal) to give 3-methylcytosine, m. p. 213°, unchanged by further recrystallisation from ethanol. [Found (in material dried at 125°): C, 47·7; H, 5·35; N, 32·8. $C_5H_7N_3O$ requires C, 48·0; H, 5·64; N, 33·6%.] Material dried at 60° was essentially the hemihydrate (Found: C, 44·6; H, 5·9; N, 31·9. $C_5H_7N_3O,\frac{1}{2}H_2O$ requires C, 44·8; H, 6·0; N, 31·4%). Equivalent weight by titration, 129 (theoretical 125). The picrate (from ethanol), recrystallised from water, had m. p. 231—240° (lit.,[14] m. p. 244—246°). Dr. Lawley has also obtained this lower-melting modification.

* We thank Dr. R. A. Y. Jones (Sheffield University) who carried out these measurements on an AEI machine.

† Brown [13] found that hydrolysis of the ester with 2·5N-sodium hydroxide gave a product contaminated with 3-methyluracil-5-carboxylic acid. Evidently the conditions used are quite critical.

[17] Neuman, Hammond, and Dougherty, *J. Amer. Chem. Soc.*, 1962, **84**, 1506.
[18] Katritzky and Waring, *Chem. and Ind.*, 1962, 695.

1,N^4,N^4-*Trimethylcytosine.*—4-Methoxy-1-methyl-2-pyrimidone [19] (2·0 g.) was heated at 180° for 7 hr. with methanolic dimethylamine solution (50 c.c., saturated at 0°). The filtered solution was evaporated to dryness (vacuum) and the residue gave 1,N^4,N^4-trimethylcytosine (1·61 g., 62%), m. p. 175—178° (from ethyl acetate, 100 ml.) (lit.,[7] m. p. 179°).

The base with N-hydrochloric acid (1·05 moles), evaporated to dryness at 25° (vacuum), gave the *hydrochloride*, prisms, m. p. 191—199° (rapid heating) (from ethyl acetate) (Found: C, 44·5; H, 6·5; N, 21·3. $C_7H_{12}ClN_3O$ requires C, 44·3; H, 6·4; N, 22·2%).

1,N^4-*Dimethylcytosine.*—This was prepared similarly. Two recrystallisations of the product from ethanol–ethyl acetate (1 : 3) gave the pure derivative (45%), m. p. 177—178° (lit.,[7] m. p. 179°). The *hydrochloride*, prepared as above, crystallised from ethanol–ethyl acetate as irregular white crystals, m. p. 215—225° (rapid heating), with sublimation to needles above 170° (Found: C, 41·0; H, 5·8; N, 23·5. $C_6H_{10}ClN_3O$ requires C, 41·0; H, 5·7; N, 23·9%).

1-*Methylcytosine hydrochloride*, prepared as above from the free base,[7] crystallised from ethanol as a white powder, m. p. 285—300° (decomp., rapid heating) with sublimation above 240° (Found: C, 37·6; H, 4·9; N, 26·1. $C_5H_8ClN_3O$ requires C, 37·2; H, 5·0; N, 26·0%).

Cytosine hydrochloride, from cytosine, crystallised from ethanol–ether–water, had m. p. 259—265° (decomp., rapid heating) (lit.,[20] m. p. 255—265°) (Found: C, 32·8; H, 4·1; N, 28·5. $C_4H_6ClN_3O$ requires C, 32·6; H, 4·1; N, 28·5%).

2-*Pyrimidone.*—2-Aminopyrimidine was hydrolysed.[15] The crude pyrimidone was extracted with boiling benzene; the residue crystallised from ethanol, and sublimed as a new polymorph, needles, m. p. 159—161° (lit., m. p. 178—180°,[15] 179—181° [16]) (Found: C, 50·2; H, 4·3; N, 29·5. Calc. for $C_4H_4N_2O$: C, 50·0; H, 4·2; N, 29·2%). Ultraviolet absorption: cation (1·17N-H_2SO_4), $\lambda_{max.}$ 309 mμ (log ε 3·77); neutral form (pH 7), $\lambda_{max.}$ 299 mμ (3·67); anion (0·01N-NaOH), $\lambda_{max.}$ 291 mμ (3·68) [lit.,[15] 309 (3·75); 299 (3·66); 290 (3·66), respectively]. When the 2-pyrimidone was crystallised from ethyl acetate, and seeded with an authentic sample provided by Dr. D. J. Brown, we obtained needles, m. p. 178—179°. Dr. Brown had also obtained the lower m. p. polymorphic form.

Nuclear magnetic resonance spectra were obtained at 40 Mc./sec. using a Perkin-Elmer permanent magnet spectrometer. Positions of peaks are quoted as chemical shifts on the τ scale and have been measured against tetramethylsilane as an internal reference, or for aqueous solutions against t-butyl alcohol as an internal reference.

We thank Dr. P. D. Lawley for a discussion, him and Dr. D. J. Brown for specimens, and D.S.I.R. for a maintenance grant (to A. J. W.).

THE UNIVERSITY CHEMICAL LABORATORY, CAMBRIDGE. [*Received, October 22nd,* 1962.]

[19] Hilbert and Johnson, *J. Amer. Chem. Soc.*, 1930, **52**, 2001.
[20] Hinman, Caron, and De Boer, *J. Amer. Chem. Soc.*, 1953, **75**, 5867.