SPECTROPHOTOMETRIC STUDIES OP NUCLEIC ACID DERIVATIVES AND RELATED COMPOUNDS AS A FUNCTION OF pH

I. PYRIMIDINES

by

DAVID SHUGAR AND JACK J. FOX*

Faculté des Sciences, Laboratoire de Morphologie Animale, Université Libre de Bruxelles (Belgium)

INTRODUCTION

A series of preliminary studies of a number of natural and synthetic pyrimidine nucleosides demonstrated that net differences in absorption relative to the carbohydrate components of these compounds do exist. In view of the possibility of the use of absorption spectra as a criterion for the characterization of the sugar moiety of these biologically important substances, this study has been extended, the details of which will be presented in a forthcoming publication.

During the course of our work on nucleoside spectra it was found necessary to account for the spectra of the pyrimidine bases themselves and in particular their variation as a continuous function of pH. The results of this latter study are reported here.

The ultraviolet absorption spectra of pyrimidines have been studied extensively, in many instances at different pH values. It is, however, most surprising that in almost all instances the pH values at which these spectra were measured were selected either at random, or only at such values as to give pronounced differences in spectra from one pH to another.

It was shown long ago by STENSTROM AND GOLDSMITH[1] that where the absorption spectrum of a compound is dependent upon the pH of the medium, the study of these changes may be used to demonstrate the limiting neutral or ionic forms of the substance as well as the dissociation constant. The principle of this method is evident from an examination of Fig. 6 which shows the spectrum of 1-methyluracil between pH 7.2 and 14. It is known from the work of LEVENE and co-workers[2] that the number 4 hydroxyl group dissociates in basic medium with a pK of 9.71 determined by electrometric titration. Examination of Fig. 6 shows that from acid pH to about pH 8.6 the spectrum of this compound is represented by Curve A. As the pH is further increased changes occur until at about pH 11.0 we obtain Curve D which remains unaltered upon further increase in alkalinity to 1*N* NaOH. Curve D represents the completely-dissociated form of the molecule. It will be noticed that curves A and D cross at 2455 A, indicating that

* Fellow of the American Cancer Society, recommended by the Committee on Growth of the National Research Council. Present address, Sloan-Kettering Institute for Cancer Research, 444 East 68th Street, New York, N.Y.

References p. 218.

both neutral and dissociated forms of the molecule have the same extinction coefficient at this particular wave-length. All curves at intermediate pH values (*i.e.*, B and C) should therefore pass through this isosbestic point, *a*, as they indeed do. Furthermore, that curve which lies exactly midway between A and D represents a mixture in which one-half of the molecules are in the dissociated form and one-half in the neutral form; so that the pH of the solution giving this curve should correspond to the pK of the compound. If a sufficient number of pH values are run, a plot of extinction values at any one wave length versus pH would give a "titration" curve of the compound. In practice, however, it is usually sufficient to run several curves at pH values in the vicinity of the pK to determine the pK by simple extrapolation to within 0.05 or 0.1 of a pH unit. For example, the value of the pK obtained from Fig. 6 is 9.75, as compared to the value of 9.71 obtained by LEVENE electrometrically.

It is apparent from the above that a knowledge of the pK of a substance is essential for an adequate interpretation of its spectrum. A clear understanding of this point is exhibited in a recent publication by MARSHALL AND WALKER[3] who first determined the pK's of their compounds electrometrically, following which they made spectral measurements at pH values sufficiently removed from these pK values so as to eliminate the possibility of the presence of more than one species of the compound in solution.

In the work outlined below, however, the spectra of the compounds under study have been investigated over a wide enough pH range to show spectrophotometrically the limiting ionic species in each case, and to permit at the same time the calculation of the pK of the compound*. It will be seen that this method is the more useful in that it provides a continuous picture of what is taking place particularly when, (1) the likelihood or possibility of the existence of more than one dissociating group is suspected, (2) the dissociation may occur at pH values above 10 or 11 where electrometric or titrimetric measurements are rendered difficult, and, in some cases, impossible, and (3) where alteration of structure in solutions of varying pH's is suspected, *viz.*, the lactam-lactim tautomerism of hydroxypyrimidines.

It is believed that a reasonably accurate knowledge of the spectra of pyrimidines and purines, (as well as other nucleic acid derivatives, the nucleosides, nucleotides, etc.,) as well as the variation of these spectra with pH and a complete explanation of this variation in terms of structure and properties, will be of value in studies of the structure and composition of nucleic acids in the same way, perhaps, that a knowledge of the spectra of aromatic amino acids has been applied to the analysis and structure of proteins[4]. Furthermore, a knowledge of the variation of the spectra of nucleic acid hydrolysis products with pH is becoming increasingly important as a result of the use of spectrophotometric methods for studying the enzymatic activity of nucleases[5].

EXPERIMENTAL

All measurements were made with a Beckman Model DU spectrophotometer, Serial No. 29543, using 10 mm quartz cells. Cell corrections were established at frequent wave-length intervals over the entire wave-length range used with distilled water. Solutions were pipetted and mixed directly in the cells, using Carlsberg constriction micropipettes and precision-graduated 1 and 2 ml pipettes. Concentrations were generally adjusted so that most of the important optical density readings were taken in the range from 0.3 to 0.8 at wave-length intervals from 5 to 20 A, the smaller interval

* No attempt has been made in this report to measure the basicity of the ring nitrogen atoms by this method.

References p. 218.

being used in the neighborhood of maxima, minima, or points of inflection. In the region from 3500 A to 2300 A slit widths varied from 0.15 to 0.25 mm. Owing to the fact that the instrument was a new one with untarnished optical surfaces, it was found possible to make measurements down to 2050 A in those instances where buffer absorption so permitted. In view of the possibility of stray light at wave-lengths below about 2200 A, the spectra of 2,4-diethoxypyrimidine (Fig. 11) and 1-methyl-4-ethoxy-2-pyrimidone (Fig. 10) were compared. These spectra should show pronounced differences (see below), and in fact it is found that the former displays a maximum at 2100 A while the latter does not. A study of various concentrations of the dialkoxy derivative showed that it obeyed BEER's Law quite accurately at 2050, 2100, and 2150 A up to an optical density of over 1.0. Furthermore, in several instances the pK value for a compound calculated at 2100 A was the same as that calculated at higher wave-length values. However, the accuracy of the measurements at these shorter wave lengths is of necessity lower owing to the use of slit widths up to 2.0 mm at the very lowest wave length used.

All results are expressed as molar extinction coefficients, defined as $\varepsilon = \frac{d}{c}$ where d is the optical density in the 10 mm cell at a concentration, c, expressed in moles per litre. It is believed that the results obtained are correct to within 0.5 to 1.0%.

MATERIALS

Buffers. 0.1 *N* HCl was taken as pH 1 and 0.01 *N* HCl as pH 2.0. Between pH 2.0 and 5.4 Walpole acetate buffers were used at a normality of 0.025. Between pH 6.0 and 8.6 Sorensen phosphate buffers were used at 0.015 *M*. Above this range Sorensen's glycocol buffers were employed. Finally, 0.01 *N* NaOH was taken as essentially equal to pH 12.0, 0.1 *N* NaOH as pH 13, 0.5 *N* and 1.0 *N* NaOH assumed to be pH 13.5 and 14 respectively. Except for the extremes, pH values were checked with the glass electrode, the latter being calibrated with several standard buffers.

Compounds. Uracil and *thymine* were products of the Schwarz Laboratories, Inc., and were recrystallized several times from hot water until constant values were obtained spectrophotometrically. 1-*Methyl-4-ethoxy-2-pyrimidone* was prepared from 2,4-diethoxypyrimidine by the method of HILBERT AND JOHNSON[6] and recrystallized to a constant melting point, 135–136°. 1-*Methyluracil*, the hydrolysis product of the latter compound, was recrystallized thrice from alcohol, m.p. 231–232°. 5-*Nitrouracil* was synthesized according to JOHNSON AND MATSUO[7] and repeatedly recrystallized from water. 1,3-*Dimethyluracil* was prepared from uracil by methylation with dimethylsulfate[8]. Three recrystallizations from alcohol-ether solution gave a pure product, m.p. 123–124°.

3,4-*Diethoxypyrimidine*[6,9] was redistilled thrice under vacuum. 2-*Ethoxy- and* 4-*ethoxyuracil* were prepared according to HILBERT AND JANSEN[10]. After repeated recrystallizations they were chromatographed separately and in admixture in propanol-ammonia-water. A clear separation was obtained for the mixture whereas the individual compounds gave but one spot each on the chromatogram. M.p. 2-ethoxyuracil, 127–128°; 4-ethoxyuracil, 166–167°. 2-*Methoxy-4-aminopyrimidine* was synthesized by the method of HILBERT AND JOHNSON[9]. In spite of repeated recrystallizations from boiling water, minute traces of the 2-amino-4-methoxy derivative could be detected by chromatograms in 5% NaH_2PO_4-*iso*amyl alcohol. The compound melted at 172–174° and was considered sufficiently pure for spectrophotometric measurements.

Cytosine was obtained from Dr G. B. BROWN of the Sloan-Kettering Institute for Cancer Research, 5-*methylcytosine* from Dr G. H. HITCHINGS of the Wellcome Research Laboratories, *Orotic Acid* from Dr H. K. MITCHELL of the California Institute of Technology, 3-*Methyluracil* from the Levene Collection through the courtesy of Dr A. E. MIRSKY of the Rockefeller Institute for Medical Research, to all of whom the authors are deeply indebted. The above 4 compounds proved to be pure both by melting points and spectrophotometric behaviour.

All of these compounds were dried at temperatures between 115–160° as their melting points permitted.

RESULTS

Cytosine and 5-*methylcytosine.* The variation of the spectrum of 1-methyluracil has been discussed above and explained in terms of ionic dissociation of the 4-hydroxyl group in alkaline media. In the case of cytosine and 5-methylcytosine two potentially-dissociable groups are present, the 4-amino and 2-hydroxy, the effects of which may be expected to manifest themselves in their spectra. Examination of the spectrum of cytosine (Fig. 1) in fact reveals the presence of two equilibria, one in acid media and the other

References p. 218.

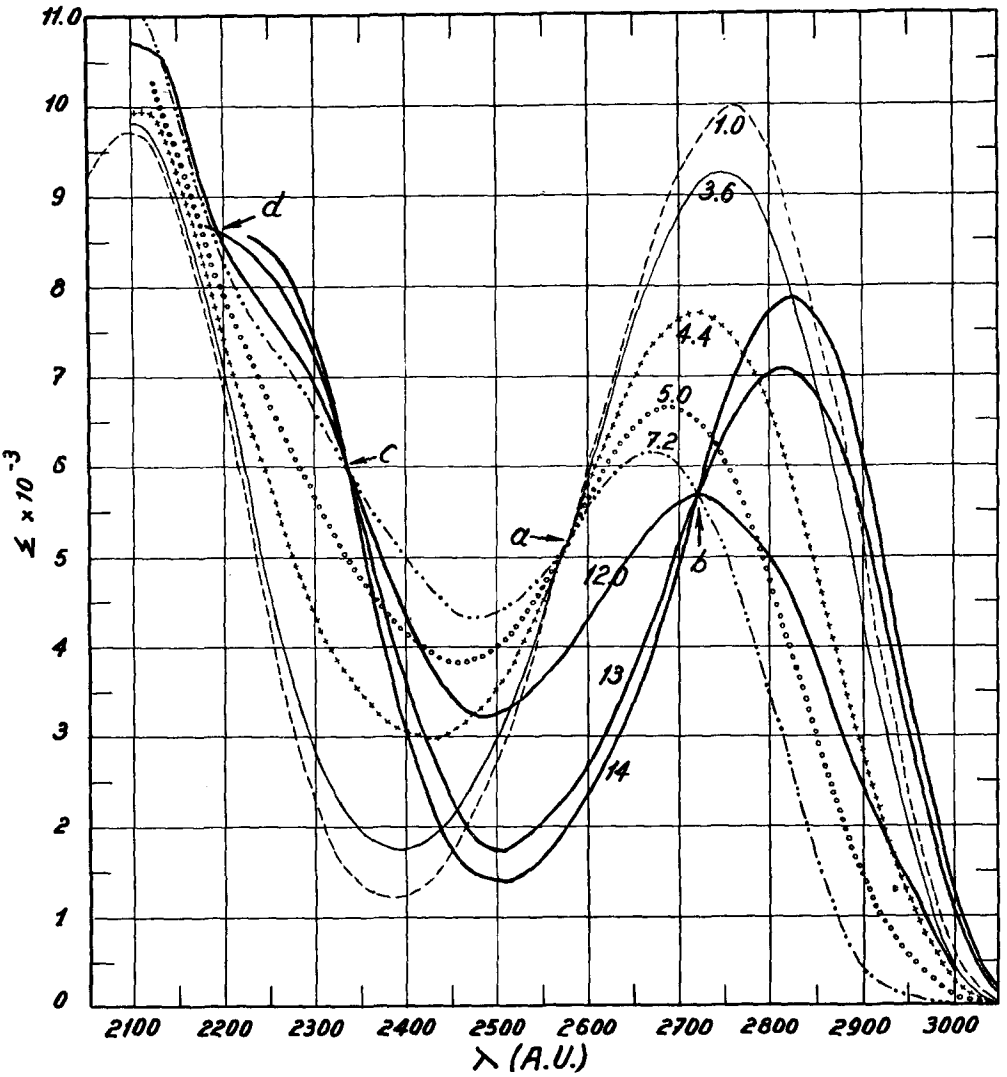


Fig. 1. *Cytosine* in aqueous solution at pH values indicated. Isosbestic point *a* is that for pK_1; *b*, *c*, and *d* for pK_2. The curve for pH 2, not shown, is identical with that for pH 1; that for pH 11.0, also absent, is only slightly displaced from pH 7.2 curve. (At pH 7.2, ε increases to $12.7 \cdot 10^3$ at 2075 A, the shortest wave-length at which measurements were made). $pK_1 = 4.45$, $pK_2 = 12.2$.

in base, each giving spectrophotometrically-determined pK values in good agreement with those determined by LEVENE and co-workers[2] by electrometric titration (see Table 1).

The discovery of 5-methylcytosine as a constituent of nucleic acid[11] has since been supported by HOTCHKISS[12] and confirmed by WYATT[13]. The latter author listed its ultraviolet absorption spectra at 3 pH's, acid, neutral, and base. The pK's of this compound, to our knowledge, have not been reported. Fig. 2 for 5-methylcytosine shows an analogous picture to that of cytosine, as one would expect, disregarding for the moment the batho-

References p. 218.

chromic shift caused by the 5-methyl group. pK values of 4.6 and 12.4 for the amino and hydroxyl group respectively are calculated from its spectrum.

TABLE I

SPECTROPHOTOMETRICALLY DETERMINED "APPARENT" DISSOCIATION CONSTANTS COMPARED TO VALUES DETERMINED BY TITRATION

Compound	*Dissociation Constant(s)*	
	Spectrophotometric	*Titrimetric*
Cytosine	4.45	4.60 (2)*
	12.2	12.16 (2)
5-Methylcytosine	4.6	—
	12.4	—
Uracil	9.5	9.45 (2)
	> 13	—
Thymine	9.9	9.94 (2)
	> 13	—
1-Methyluracil	9.75	9.71 (2)
3-Methyluracil	9.95	9.99 (2)
1,3-Dimethyluracil	none	none (2)
2-ethoxy-4-hydroxypyrimidine	8.2	8.4** (3)
4-ethoxy-2-pyrimidone	10.7	—
5-Nitrouracil	5.3	5.48 (24)
	11.7	—
Orotic Acid	~ 2.8	2.40 (20)
	9.45	—
	> 13	—
2-methoxy-4-aminopyrimidine	5.3	—

* Numbers in parentheses are literature references.
** This value for 2-methoxy-4-hydroxy-6-methylpyrimidine.

Uracil. The dissociation of uracil in basic solution has been the subject of several studies. In 1925, LEVENE AND SIMMS[14] listed two constants for uracil, pK 9.28 and 13.56, which they determined electrometrically. Later, LEVENE and co-workers[2] discarded the higher constant on the basis of experimental errors in the activity corrections and listed uracil as a weak acid with but one pK at 9.45. They concluded that positions 2 and 4 have equal tendencies to "enolize and ionize" since the pK's of 1- and 3-methyluracil gave values of 9.99 and 9.71 respectively. They considered the enolization of one position to inhibit the enolization of the other thus giving but one dissociation constant for uracil.

Fig. 3 shows the spectrum of uracil. Close examination reveals that between neutrality and pH 10.5 all curves pass through isosbestics *a* and *b*. Calculation of the pK in this range would give a value of 9.3 in close agreement with the first pK value listed by LEVENE[14]. The curves between pH 11.0 and 0.1 *N* NaOH are identical but slightly displaced (about 3%) from isosbestic point *a*. The curves for higher pH (above 0.1 *N* NaOH) show decided shifts to the shorter wave-lengths and concomitantly possess new isosbestic points *c*, *d*, *e*, and possibly *f*, denoting a new equilibrium with a pK_2 greater than 13.

Because of evidence which will be presented later, we feel inclined to include in the first ionic equilibrium all curves between neutrality and pH 12.0 which gives a pK value of 9.5 in agreement with LEVENE's redetermined value for uracil[2].

References p. 218.

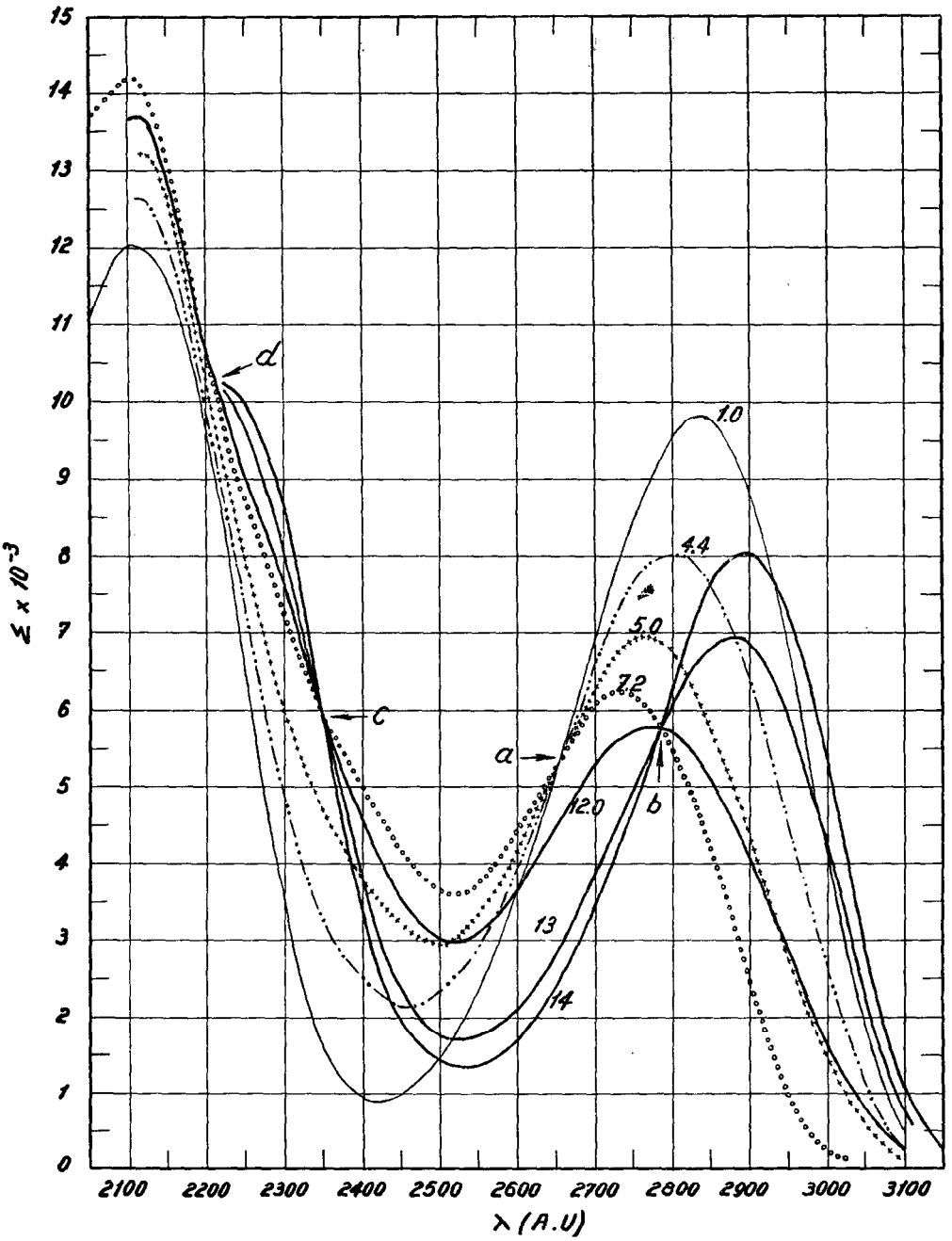


Fig. 2. *5-Methylcytosine* in aqueous solution at pH values indicated. Isosbestic point *a* is that for pK_1; *b*, *c*, and *d* for pK_2. The curve for pH 2.0, not shown, is identical with pH 1 curve. $pK_1 = 4.6$, $pK_2 = 12.4$.

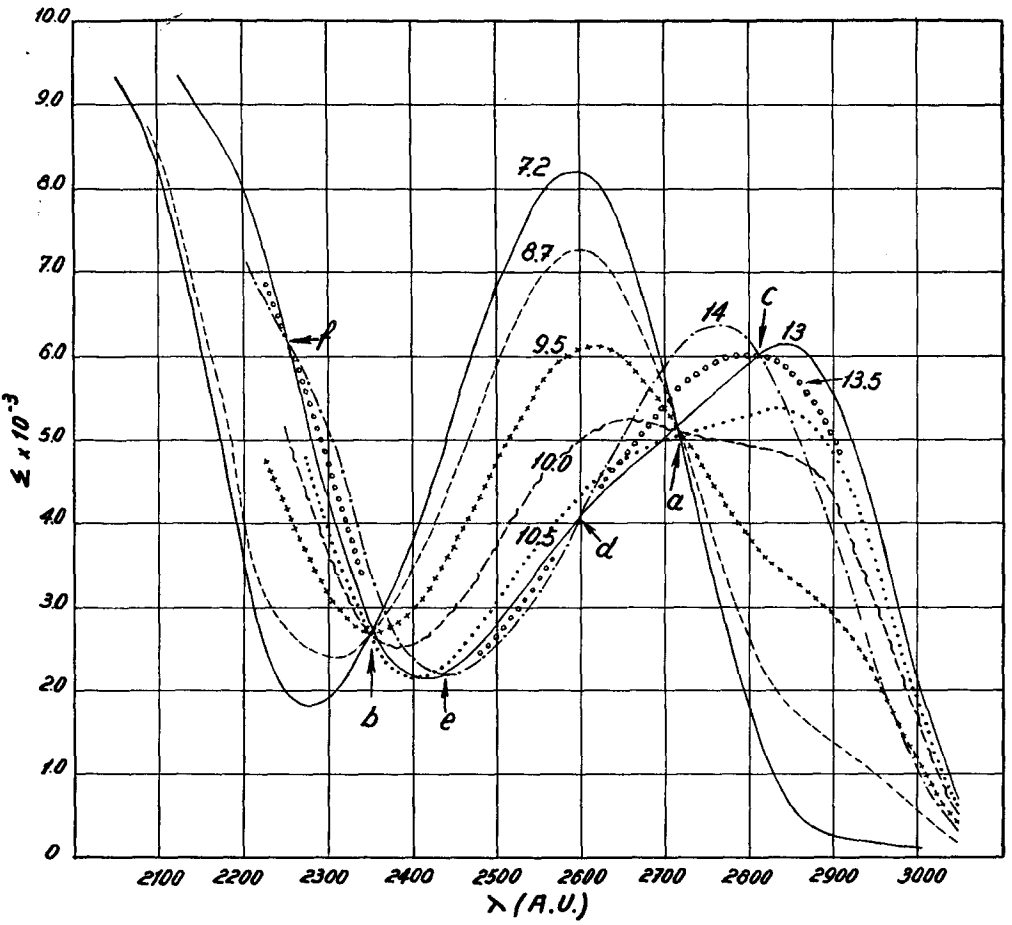


Fig. 3. *Uracil* in aqueous solution at pH values indicated. Isosbestic points *a* and *b* are those for pK_1; *c*, *d*, *e*, and *f* for pK_2. Curves for pH 11.0 and 12 are practically identical with that for pH 13. $pK_1 = 9.5$, $pK_2 > 13$.

Regarding the curves between pH 7.2 and 11.0, there can be little doubt as to their validity since the values for the extinction coefficients and the position of the maxima and minima are in very close agreement with those obtained by PLOESER AND LORING[15] on synthetic as well as natural samples of uracil at these pH's. Further, for the curve at pH 14, acidification of the sample in the quartz cell gives a curve identical with that for pH 7.2, thus confirming the reversibility of both equilibria.

Thymine. A similar picture should result from the spectrum of thymine (Fig. 4) and indeed, apart from the bathochromic shift due to the 5-methyl substituent, this is the case. The pK_1 determined spectrophotometrically is in excellent agreement with LEVENE'S value (Table 1). Above pH 13 a new equilibrium due to the dienolic form is in evidence. The pK_2 for the latter equilibrium is well above 13 and is higher than the corresponding pK_2 value for uracil. This is concluded on the basis that the pH 14 curve is less displaced from the pH 13 curve in thymine than are the corresponding curves for uracil. The slight deviation from isosbestic point *a*, noted for uracil, is not present for

References p. 218.

TABLE II

SPECTROPHOTOMETRIC DATA OF PYRIMIDINE DERIVATIVES

Compound	*Absorption maxima**			*Isosbestic points*		
	pH	λ (*A*)	$\varepsilon \cdot 10^{-3}$	*pH range*	λ (*A*)	$\varepsilon \cdot 10^{-3}$
Cytosine	1.0–2.0	2100, 2760	9.70, 10.0	1–10	2580	5.20
	3.6	2745	9.25			
	4.4	2720	7.73	10–14	2195	8.6
	5.0	2690	6.65		2335	5.90
	7–10	2670	6.13		2720	5.63
	12.0	2720	5.63			
	13	2815	7.06			
	14	2820	7.86			
5-Methyl-cytosine	1.0–2.0	2105, 2835	12.0, 9.79	1–10	2660	5.40
	4.4	2105, 2810	12.6, 8.00			
	5.0	2105, 2770	13.2, 6.94	10–14	2210	10.4
	7–10	2105, 2735	14.2, 6.23		2345	6.05
	12.0	2775	5.80		2790	5.77
	13	2880	6.95			
	14	2895	8.05			
Uracil	4.4–7.2	2595	8.20	4–10.5	2350	2.68
	8.7	2600	7.30		2720	5.08
	9.5	2615	6.11			
	10.0	2660	5.25	13–14	2255	6.2
	10.5	2840	5.40		2440	2.2
	12–13	2840	6.15		2600	4.07
	13.5	2805	6.05		2810	6.02
	14	2765	6.38			
Thymine	4.4–7.2	2070, 2645	9.5, 7.89	4–13	2395	2.52
	10.0	2670	5.97		2780	5.02
	12–13	2910	5.44			
	13.5	2880	5.30	13–14	2285	6.4
	14	2820	5.45		2460	2.17
					2640	4.12
					2875	5.30
1-Methyl-uracil	5.4–7.2	2075, 2675	8.8, 9.75	3–14	2455	3.66
	9.5	2665	8.78			
	10.0	2655	7.95			
	12–14	2650	7.02			
3-Methyl-uracil	3.0–7.2	2585	7.30	3–14	2350	2.43
	9.5	2615	6.22		2680	5.73
	10.1	2820	6.58			
	12–14	2180, 2825	7.06, 10.7			
1,3-Dimethyl-uracil	1–14**	2660	8.90	—	—	—
5-Nitro-uracil	1.0–3.0	2375, 3000	7.18, 9.45	1–10	2220	5.4
	5.0	2360, 3055	6.70, 7.82		2490	5.0
	5.6	2330, 2600, 3405	6.37, 4.80, 11.5		2680	4.0
					3140	7.75
	7.2–10.0	2305, 2580, 3420	6.0, 5.50, 15.7	10–14	2550	5.35
	12.0	2425, 3550	6.2, 14.5		2720	2.64
	13	2425, 3620	6.3, 15.6		2970	3.12
	14	2425, 3620	6.4, 15.8		3515	14.3

TABLE II (Continued)

Compound	Absorption maxima*			Isosbestic points		
	pH	λ (A)	ε·10⁻³	pH range	λ (A)	ε·10⁻³
Orotic Acid	1.0–2.0	2050, 2800	10.9, 7.52	1–7	2830	7.43
	4.4–7.2	2070, 2785	11.6, 7.68			
	9.5	2815	6.58	7–13	2515	2.75
	12–13	2860	5.98		2920	5.60
	13.5	2865	5.68			
	14	2890	5.35	13–14	2415	2.6
					2950	5.13
2-ethoxy-4-hydroxy-pyrimidine	4.4–7.2	2175, 2595	6.65, 6.05	4–13	2335	3.45
	8.4	2640	6.32		2595	6.05
	10–13	2205, 2645	7.03, 6.70		2755	4.07
4-ethoxy-2-pyrimidone	4.4–7.2	2690	5.08	4–14	2410	1.32
	10.1	2705	5.13		2690	5.08
	10.5	2740	5.28			
	13	2200, 2780	9.7, 6.43			
	14	2780	6.53			
1-Methyl-4-ethoxy-2-pyrimidone	7.2	2745	6.10	—	—	—
2-methoxy-4-amino-pyrimidine	2.0	2295, 2605	8.82, 9.47	1–7	2180	6.6
	7.2	2250, 2705	8.0, 7.20		2235	7.9
					2695	7.15
2,4-diethoxy-pyrimidine	7.2	2105, 2590	7.30, 6.12	—	—	—

* Points of inflexion are not given.
** See text for behaviour of this compound in 1 *N* NaOH.

thymine. It may be noted that at neutral pH the extinction for uracil and thymine are within 3% of each other while the position of the maximum in thymine is displaced 50 A toward the longer wave lengths. A similar situation is encountered with cytosine and 5-methylcytosine. Though we have not examined the spectrum of 6-methyluracil, it is reasonable to assume that it, too, will exhibit two ionic equilibria in basic solution.

The existence of a second dissociation indicated by the spectrum of uracil may bear important structural implications. The fact that these results are in contradiction with those of LEVENE and co-workers[2] and the importance of accurate and dependable measurements of the absorption spectra of nucleosides and nucleotides as well as their free bases in alkaline media, prompted us to undertake a study of a series of pyrimidine derivatives in which 0, 1, 2 and 3 potentially dissociable groups are present.*

1,3-*Dimethyluracil.* As would be expected, this compound, fixed in the diketonic form, has no dissociable groups, a fact shown titrimetrically by LEVENE[2]. The spectrum of this compound in alcoholic solution has been reported by AUSTIN[16] and in aqueous solution by LOOFBOUROW and co-workers[17]. The latter authors show a variation in the spectrum of 1,3-dimethyluracil with pH, the neutral curve of which differs from those

* Barbituric Acid, which exhibits two equilibria spectrophotometrically, is discussed along with several of its derivatives in a separate paper[27].

References p. 218.

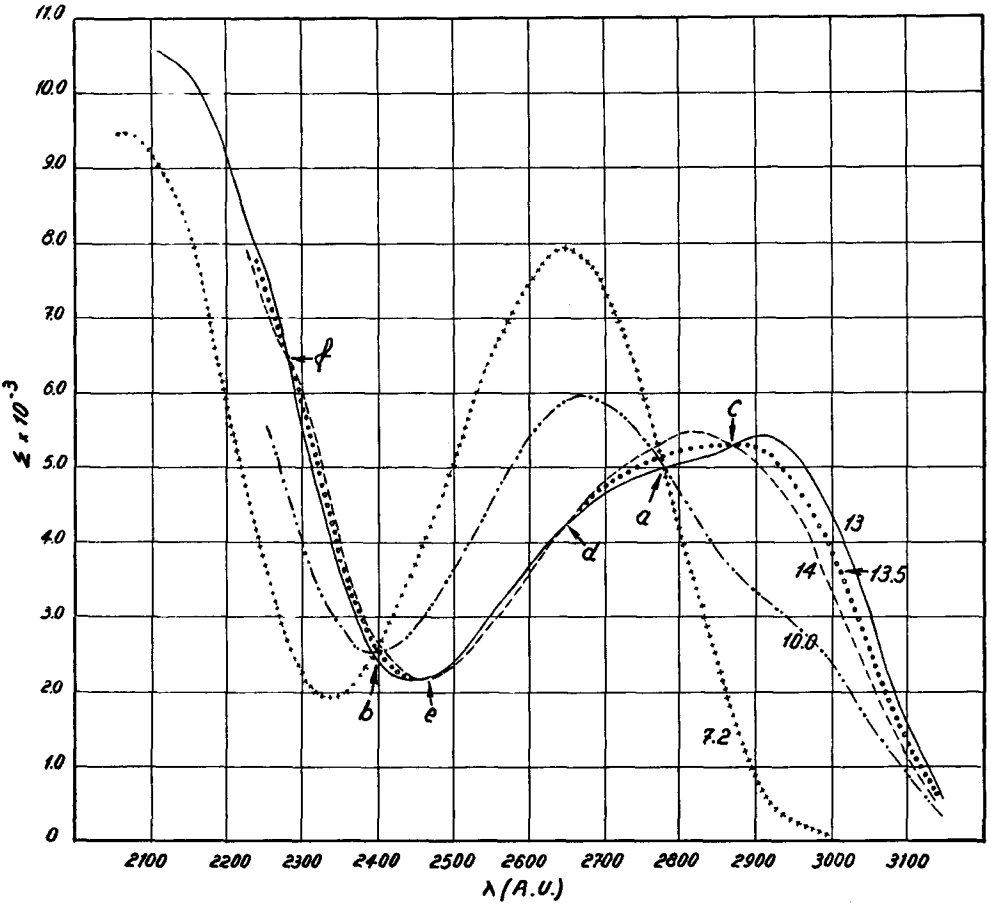


Fig. 4. *Thymine* in aqueous solution at pH values indicated. Isosbestic points *a* and *b* are those for pK_1; *c*, *d*, *e*, and *f* for pK_2. Curve for pH 12.0, not shown, is identical with that for pH 13. $pK_1 = 9.9$, $pK_2 > 13$.

in acid and alkaline solution. The curve we obtain (Fig. 12) shows no variation from pH 1 up to 14 in full accord with the absence of dissociable groups.

It was, however, noted that at 1*N* NaOH the extinction coefficient of this compound increased *slowly with time* over the entire spectral range*. Neutralization of the solution in the quartz cells did not give the original curve but rather one with extinction coefficients somewhat lower, from which it is evident that some chemical reaction occurs in 1*N* NaOH. In order to verify this conclusion, a concentrated solution of 1,3-dimethyluracil was allowed to stand overnight in 1 *N* base after which it was neutralized and chromatographed using two solvent systems, propanol-ammonia-water and butanol-water. Both chromatograms showed only a small quantity of 1,3-dimethyluracil in the treated samples along with several other components of widely varying R_F values. Untreated dimethyluracil showed only one spot on these chromatograms. No further

* Of all the compounds studied in this report, 1,3-dimethyluracil is the only instance where a change in spectrum as a function of time was noted.

References p. 218.

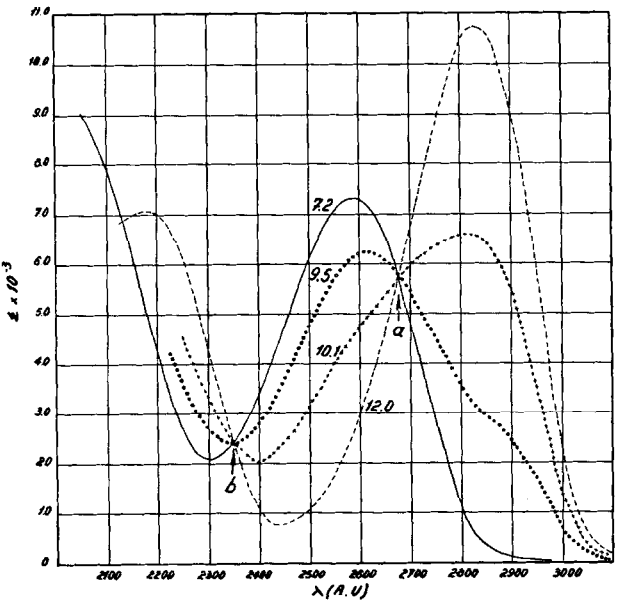


Fig. 5. *3-Methyluracil* in aqueous solution at pH values indicated. Curves for pH 11.0, 13, and 14, not indicated, are identical with that for pH 12.0. pK = 9.95.

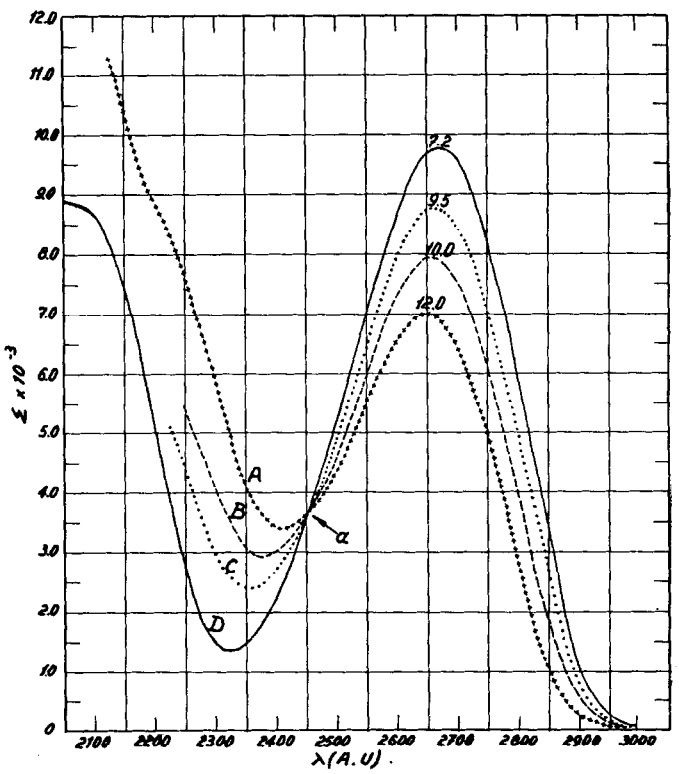


Fig. 6. *1-Methyluracil* in aqueous solution at pH values indicated. Curves for pH 11.0, 13, and 14, not shown, coincide with that for pH 12.0. pK = 9.75.

attempt was made to identify the nature of the products resulting from this rather unexpected reaction.

References p. 218.

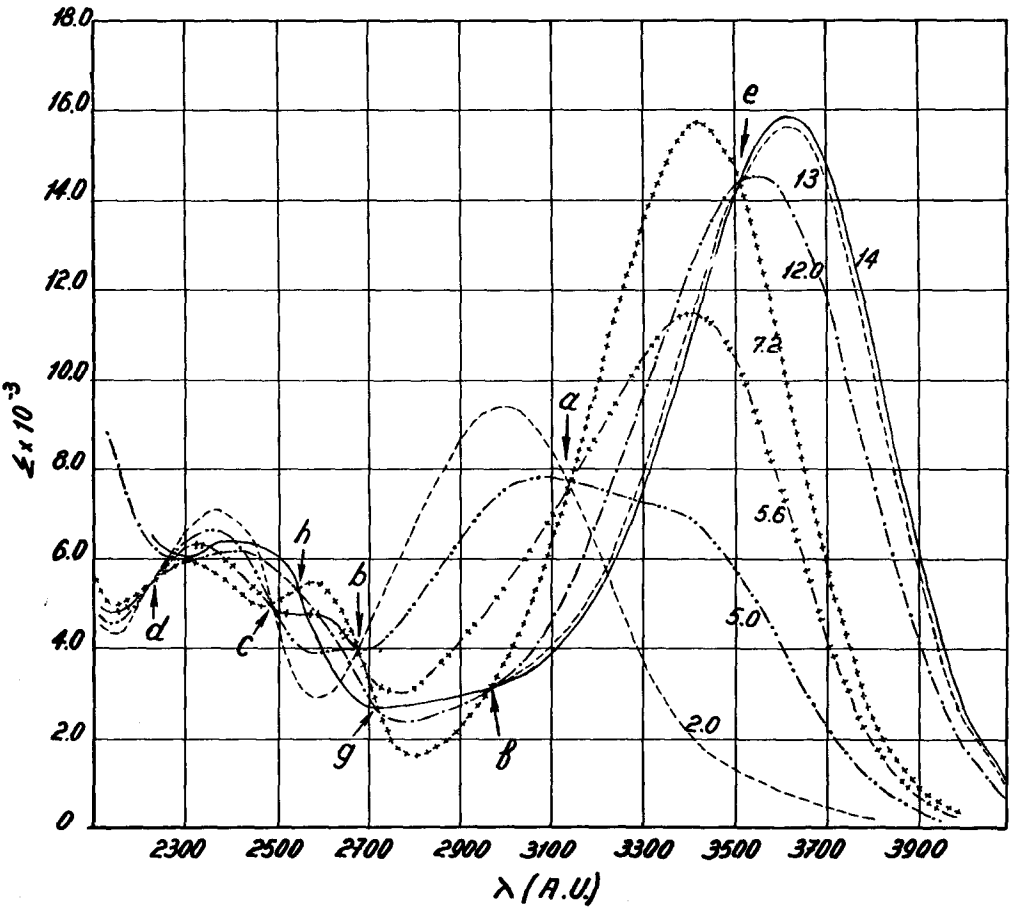


Fig. 7. 5-*Nitrouracil* in aqueous solution at pH values indicated. Isosbestic points *a*, *b*, *c*, and *d* correspond to pK_1; *e*, *f*, *g* and *h* are those for pK_2. The curves for pH 12.0 and 14 are not shown below 3000 A, while that for pH 11.0, necessary for the calculation of pK_2, has been omitted for the sake of clarity. Its ε at 3700 A is $8.3 \cdot 10^3$. $pK_1 = 5.3$, $pK_2 = 11.7$.

3-Methyluracil (Fig. 5) and 1-*Methyluracil* (Fig. 6) give but one ionic equilibrium as would be expected with pK values in accord with those reported in the literature.

5-*Nitrouracil*. The 5-nitro group, aside from exerting a large bathochromic shift toward the near ultraviolet[18] would be expected to increase the lability of the potentially-dissociable hydrogen atoms as a result of its strong inductive influence. Thus the dissociation picture of the second equilibrium should be manifested at lower pH's than was found to be the case with uracil and thymine. Fig. 7 shows the spectrum of 5-nitrouracil from which two equilibria are evident. The higher pK, 11.7, hitherto not reported, further confirms the validity of the higher dissociation picture of uracil.

Orotic Acid (uracil-6-carboxylic acid). This pyrimidine derivative has in recent years acquired marked biological importance. Of interest is the recent isolation of orotidine[19], a nucleoside containing the above pyrimidine in its structure. Though three potentially-dissociable groups are indicated by the structure of orotic acid, only one, to our knowledge, has been reported[20], obtained by conductometric titration. Examination of

References p. 218.

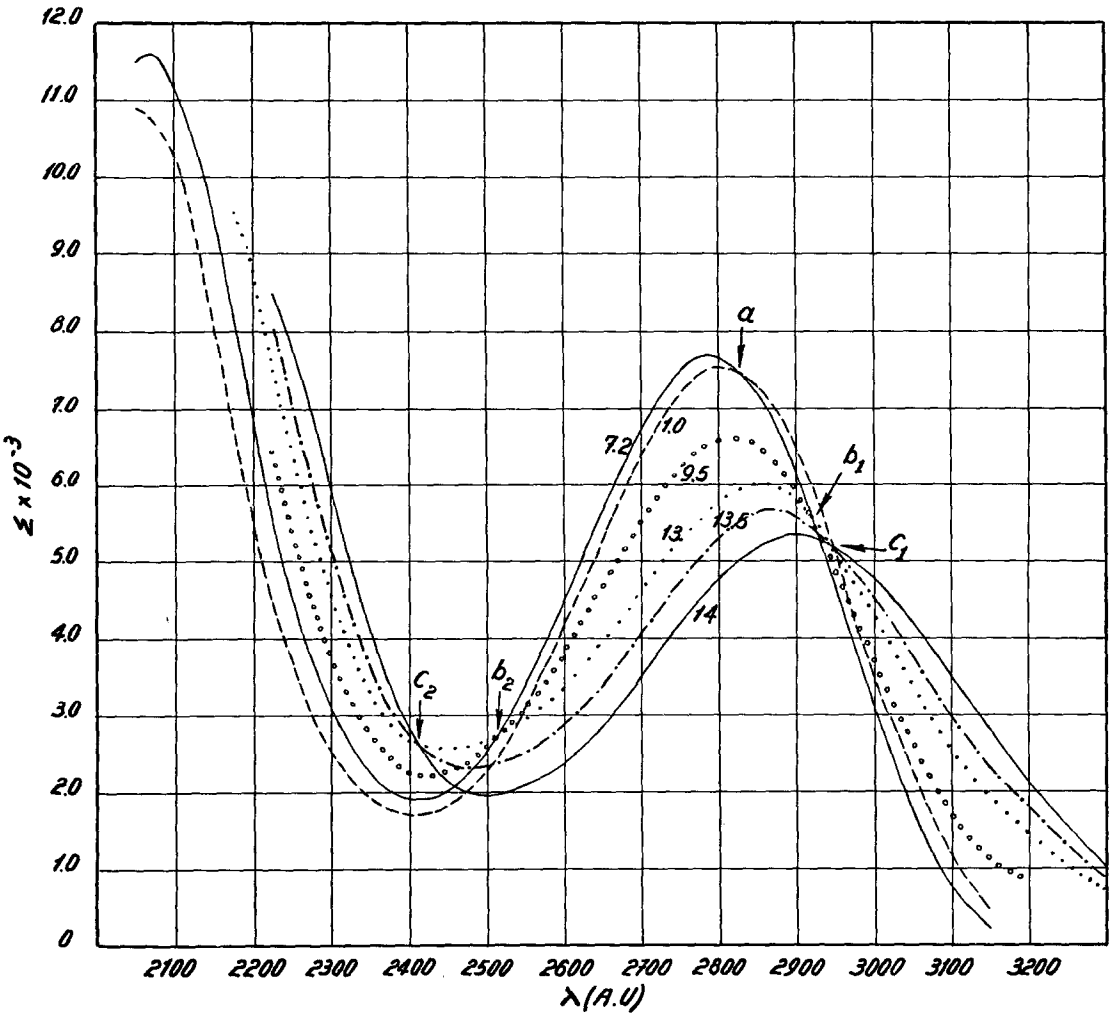


Fig. 8. *Orotic Acid* in aqueous solution at pH values indicated. Isosbestic point *a* is that for pK_1; b_1 and b_2 for pK_2; c_1 and c_2 for pK_3. The curve for pH 12.0, not shown, is identical with that for pH 13; that for pH 11.0 is only slightly displaced from that for pH 13. The curve for pH 7.2 also represents that for pH 3.6; while that for pH 1 is identical with that for pH 2.0. $pK_1 \sim 2.8$, $pK_2 = 9.45$, $pK_3 > 13$.

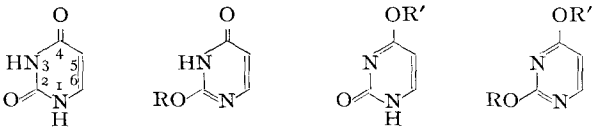
Fig. 8 shows three equilibria, one in acid and two in base. Of interest is the fact that despite the considerable alteration of the form of the curves for orotic acid from those of uracil, the pK_2 and pK_3 values are similar to those of uracil.

DISCUSSION

Thus it can be seen that spectral variation as a function of pH gives a composite picture of each of these compounds in terms of their degree(s) of dissociation and ionization constant(s). From these spectra, and others to be shown, specific information can be deduced regarding the structure of uracil in aqueous solutions of various pH values with a greater degree of certainty.

References p. 218.

Ultraviolet absorption spectroscopy has been used extensively to approach the problem of the structure of uracil*[21]. Four classical structures may be written for uracil, each a tautomeric isomer (R = R′ = H). AUSTIN[16] compared the absorption spectra of uracil in alcoholic solution to several of its derivatives. She found that uracil did not



(I) (II) (III) (IV)

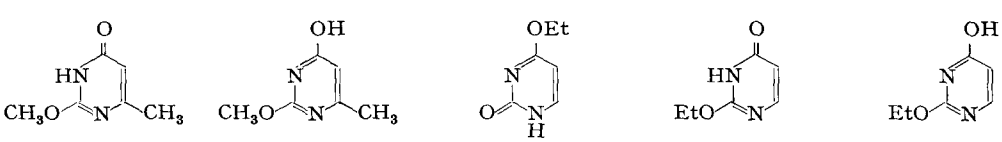
behave like 2,4-diethoxypyrimidine (IV, R = R′ = OEt), nor like 1-methyl-4-methoxy-2-pyrimidone (III, R′ = OEt). Uracil differed also from 1-methyl- and 1,3-dimethyluracil, the latter compounds giving similar spectra. Finally, since uracil gave an identical spectrum to that of 3-methyluracil in which the 4-position is ketonic, she concluded that in alcoholic solution the structure of uracil is to be represented by structure II (R = H).

During the course of our investigations, a report appeared by MARSHALL AND WALKER[3] on this problem. They consider the small bathochromic effects observed by AUSTIN in the spectra of 1-methyl-and 1,3-dimethyluracil as compared to uracil and 3-methyluracil to be a consequence of N-methylation in the same relative 1:4 position to a carbonyl or potential group analogous to similar shifts observed by SPECKER AND GAWRASCH[22] with 4-pyridone and N-methyl-4-pyridone.

Using derivatives of 6-methyluracil and taking into account the respective pK's of compounds in the choice of comparison curves, these authors show that 6-methyluracil (pK, 9.6, curve for pH 4.7) and 1, 3, 6-trimethyluracil (pH 7) give similar spectra both of which are different from that of 2-methoxy-6-methyl-4-pyrimidone (pK, 8.4, curve for pH 4.7). Hence they conclude that 6-methyluracil, and thereby uracil, is to be represented by structure I, the diketo form.

However, evidence is not presented in their report to indicate that structure (A) assigned by them to their reference compound is valid. Unless the possibility of structure (B) is eliminated, their comparisons may well have been made with a compound comparable to structure IV for uracil. The resulting dissimilarity which they observed between 6-methyluracil and its trimethyl derivative as versus their reference compound (A or B) may be comparable, in effect, to the dissimilarity observed by AUSTIN between uracil and 2,4-diethoxypyrimidine.

We were confronted with this identical problem in the course of our work using



(A) (B) (C) (D) (E)

2- and 4-alkoxy derivatives of uracil. Fig. 9 shows the spectral variation of 4-ethoxyuracil (C) with a calculated pK of 10.7. It is evident from these curves that in neutral solution this compound is in the 2-keto-4-enol form comparable to structure III for uracil (R′ = H); since at pH 13 it resembles 2,4-diethoxypyrimidine (Fig. 11) while at

* A fairly extensive review of the literature on this problem may be found in reference 3.

References p. 218.

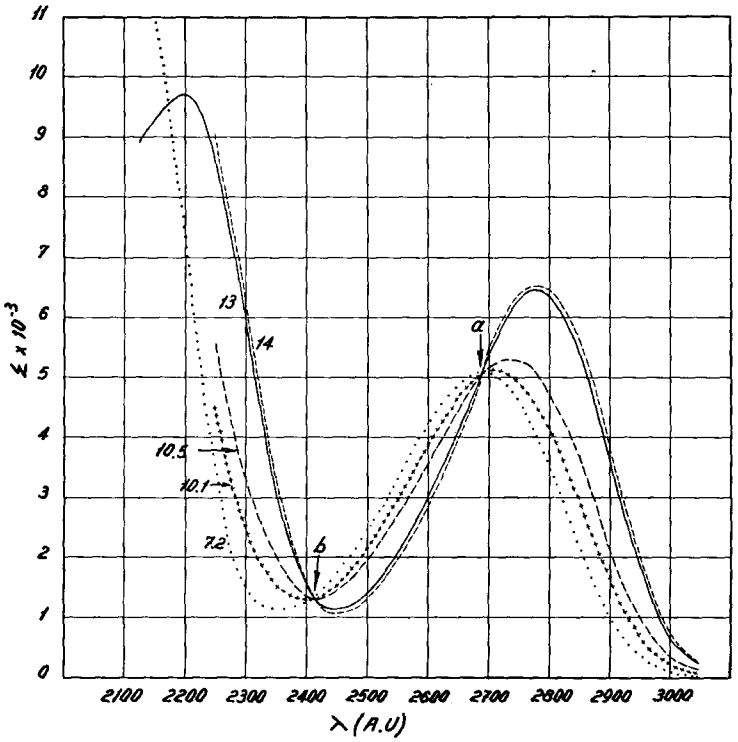


Fig. 9. 4-Ethoxy-2-pyrimidone (4-ethoxyuracil) in aqueous solution at pH values indicated. pK = 10.7.

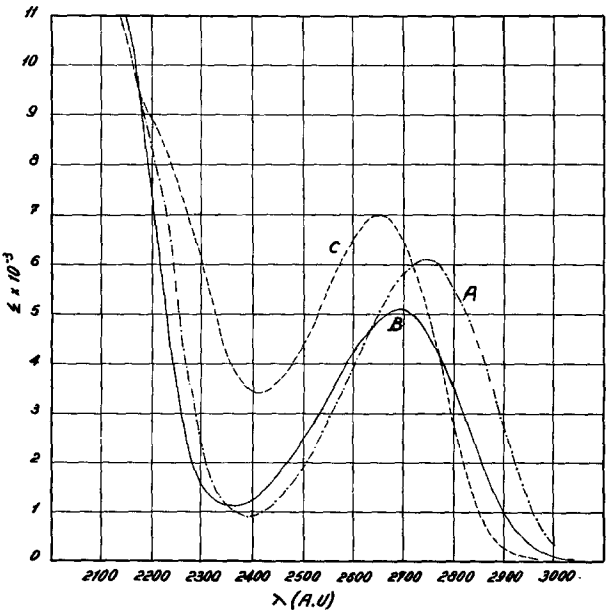


Fig. 10. *1-Methyl-4-ethoxy-2-pyrimidone* (A) at pH 7.2, *4-ethoxy-2-pyrimidone* (B) at pH 7.2, and *1-Methyluracil* (C) at pH 12.0, all in aqueous solution.

pH 7.2 it does not. Further, inspection of Fig. 10 shows that it bears close similarity to 1-methyluracil at pH 12, and to 1-methyl-4-ethoxy-2-pyrimidone.

2-Ethoxyuracil, on the other hand, gives quite a different picture, its spectrum along

References p. 218.

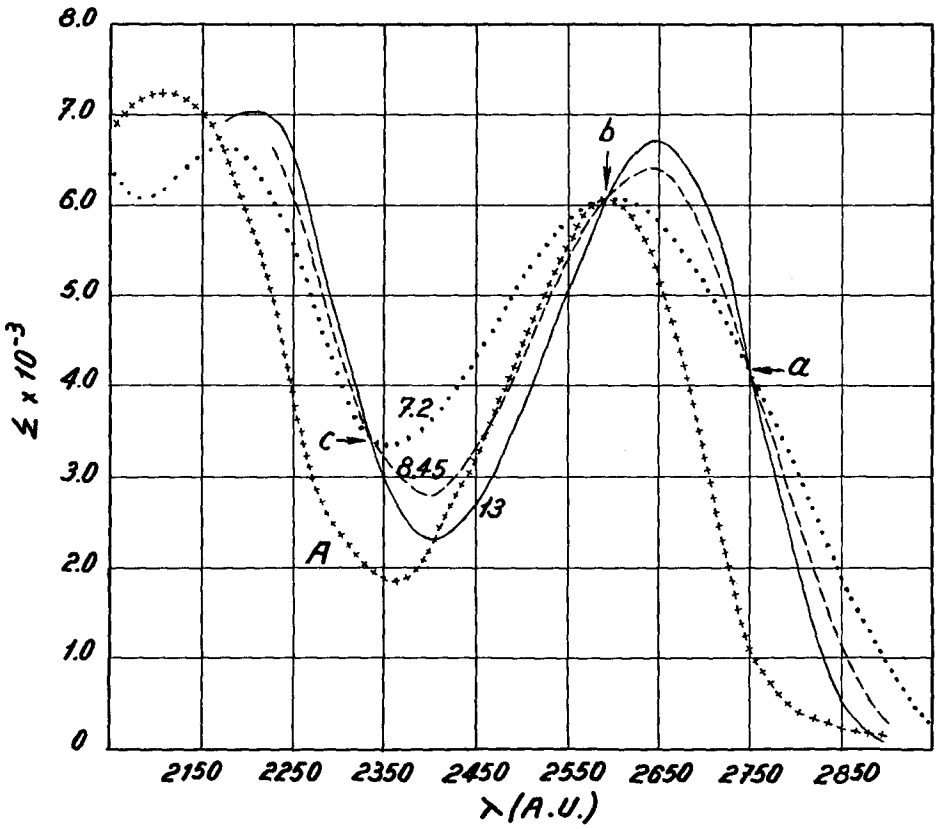


Fig. 11. *2-Ethoxy-4-hydroxypyrimidine* in aqueous solution at pH values indicated. The curve for pH 13 represents also those for pH 11.0 and 12.0; the curve for pH 4.4, not shown, is identical with that for pH 7.2. pK 8.2. Curve (A) is that for *2,4-diethoxypyrimidine* at pH 7.2.

with that of 2,4-diethoxypyrimidine being listed in Fig. 11. The pK determined spectrophotometrically is 8.2 in line with the pK determined titrimetrically by MARSHALL AND WALKER[3] for 2-methoxy-6-methyluracil (A or B). Examination of their curves for (A or B) reveals that they are similar to those for (D or E), the 6-methyl group having exerted essentially an additive effect. Comparison of the spectra of 2-ethoxyuracil with 2,4-diethoxypyrimidine shows a close resemblance regardless of the pH curve chosen for the monoalkoxy derivative. Thus, 2-ethoxyuracil is best represented by structure (E), 2-ethoxy-4-hydroxypyrimidine. On the same basis we believe the same to hold true for 2-methoxy-6-methyluracil, the latter to be represented by structure (B). We would conclude that neither of these "reference compounds" are truly representative of structure II, and as a result they neither sustain nor reject structures I or II for uracil.

A similar picture is obtained by an examination of spectral data given by STUCKEY[23] dealing with derivatives of 2-thio-6-methyluracil (F). His curve for 2-methylmercapto-6-methyluracil (G or H), taken at pH 2 (pK = 7.9) is quite dissimilar to that of 2-methylmercapto-3,6-dimethyl-4-pyrimidone (J). It may also be seen from these curves that neither of these thiouracil derivatives bear resemblance to 2-thio-6-methyluracil (F) at pH 2 or 7 (pK = 8.5). Thus compound (G or H) which is analogous to (A or B) or (D or

References p. 218.

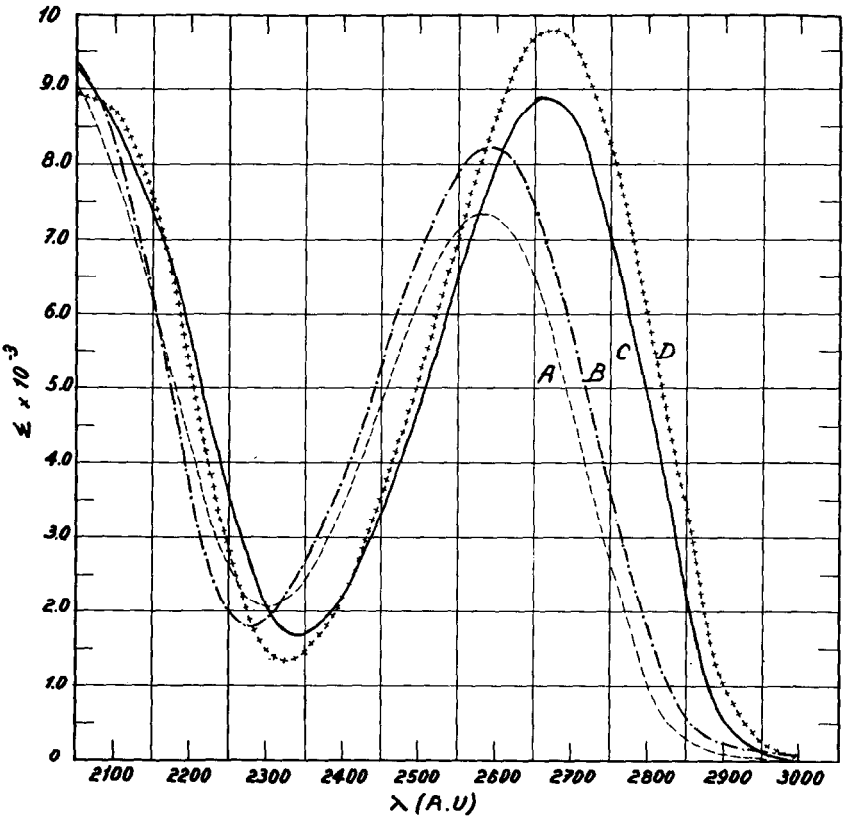
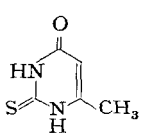
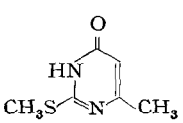
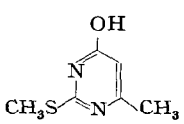
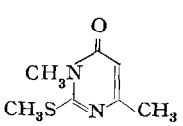


Fig. 12. *3-Methyluracil* (A), *Uracil* (B), *1,3-Dimethyluracil* (C), and *1-Methyluracil* (D), all in aqueous solution at pH 7.2.


(F)


(G)


(H)


(J)

E) is to be represented probably by (H) since it differs from the true reference compound (J). In addition, further confirmation for the diketonic structure of 2-thiouracil is provided from this comparison in line with the dipole moment measurements of SCHNEIDER AND HALVERSTADT[24].

Coming back to an examination of AUSTIN's spectral data for uracil and its N-methylated derivatives, the suggestion by MARSHALL AND WALKER that the differences observed by AUSTIN are due to N-methylation in the 1-position relative to the 4-carbonyl group is not entirely evident. Close inspection shows that uracil and 3-methyluracil exhibit the same extinction coefficient, whereas 1-methyl- and 1,3-dimethyluracil differ from each other, a fact recognized by MARSHALL AND WALKER. Furthermore, below 2400 A these resemblances between uracil and 3-methyluracil on the one hand and between 1-methyl- and 1,3-dimethyluracil on the other are not at all apparent. In view of the

References p. 218.

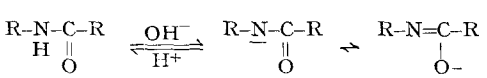
uncertainty attached to earlier spectrophotometric measurements by spectrographic methods, and the fact that AUSTIN's spectral data were determined in alcoholic media, a re-examination of these curves in aqueous solutions is warranted. Fig. 12 lists curves at pH 7 for uracil and its N-methylated derivatives as compiled from spectra already described above. It is evident that the addition of a 1-methyl group to uracil and 3-methyluracil results in identical bathochromic shifts accompanied by correspondingly identical changes in extinction coefficients. An overall inspection of these curves shows that the differences among them are due entirely to the additive effects of *both* N_1 and N_3 methylations. The inescapable conclusion is that all four compounds in Fig. 12 are in the diketo form in neutral, aqueous solutions, confirming structure I for uracil.

The structure of uracil in alkaline solutions

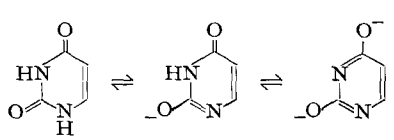
The spectrum of 3-methyluracil as a function of pH (Fig. 5), the variation of which is due to the functional group at position 2, shows a pattern in which with increasing pH the peak originally at 2585 A undergoes simultaneously a bathochromic shift and a decrease in height accompanied by the appearance of a new peak at 2835 A, the extinction coefficient of which increases. A similar pattern manifests itself for uracil (and thymine) contained in the first ionic equilibrium. While, as previously mentioned, the E_{max} of the pH 11–13 curve is not as high as that for 3-methyluracil, (in which the effect of the 4-enolic substituent is absent), the trend is the same. It is for this reason that the pH 11–13 curves, though they deviate slightly from isosbestic point *a*, are included in the first ionic equilibrium for uracil.

The characteristic pattern of 2-hydroxy dissociation is also evident for cytosine and 5-methylcytosine (Figs. 1 and 2) as well as for 5-nitrouracil (Fig. 7). PLOESER AND LORING[15] noted that the position of the long-wave maximum for uracil and cytosine were similar in base. It is noteworthy that the long-wave maximum of all curves for 3-methyluracil, cytosine, and uracil above pH 9.5 occurs at about the same wave length, the latter compound deviating only above pH 13 where a second equilibrium is manifested. It is therefore evident that the order of dissociation for uracil (and thymine) can be established from these spectra, *viz.*, the equilibrium denoted by pK_1 deals with the 2-hydroxyl while that for pK_2 (over 13) represents the dissociation of the 4-hydroxyl. It is not certain from our data whether the 4-substituent is completely dissociated in 1*N* base. A similar order of dissociation can be established for 5-nitrouracil.

With the mechanism of lactam-lactim tautomerism represented as



it is reasonable to assume that the dissociation of the lactim form at varying pH is, in effect, a measure of the lactam-lactim tautomerism itself. Thus the following order of dissociation is presented:



< pH 8.6 pH 13 > pH 13

References p. 218.

The structure of cytosine and 5-methylcytosine

The structure of cytosine in neutral solution may be inferred by comparison of its spectrum with that of 2-methoxy-4-aminopyrimidine. Fig. 13 shows the spectrum of the latter compound at pH 1 and 7.2 along with cytosine at pH 7.2 and 14. It will be seen that cytosine at pH 7.2 ($pK_2 = 12.2$) differs markedly from 2-methoxy-4-aminopyrimidine at the same pH ($pK = 5.3$); whereas, at pH 14, cytosine resembles the 2-methoxy derivative at pH 7.2. 1-Methylcytosine, the spectrum of which will be shown in a subsequent publication[26] in connection with pyrimidine nucleosides, shows an identical spectrum to that of cytosine, both at pH 7.2. It must be concluded, therefore, that cytosine, and thereby 5-methylcytosine, in solutions up to pH 10, is to be represented by the lactam form.

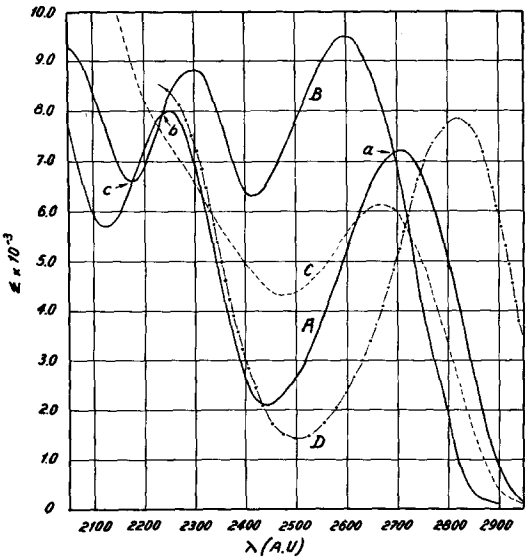


Fig. 13. *2-Methoxy-4-aminopyrimidine* at pH 7.2 (Curve A) and at pH 1.0 (Curve B). Curves for intermediate pH values, not shown, pass through isosbestic points *a*, *b*, and *c*, with a pK of 5.3. *Cytosine* at pH 7.2 (Curve C) and pH 14 (Curve D). At the shorter wave lengths the ε of curve C increases to $12.7 \cdot 10^3$ at 2075 A, while Curve D, by virtue of isosbestic point *d* of Fig. 1, passes through a maximum at about 2200 A.

SUMMARY

1. The ultraviolet absorption spectra of a number of nucleic acid derivatives (pyrimidines) and related compounds have been measured at various pH values.

2. The variation of the spectra with pH is shown in all cases to be explicable on the basis of ionic dissociation.

3. The apparent dissociation constants have been calculated from the spectra.

4. The structure of uracil (and thymine) is shown to be in the diketo form in neutral solution. Two ionic equilibria are demonstrated in alkaline solution and the order of dissociation is shown to proceed through the 2- and 4-hydroxyl groups respectively.

5. The spectra and structures of other pyrimidine derivatives are discussed.

RÉSUMÉ

1. Les spectres d'absorption dans l'ultra-violet d'un certain nombre de dérivés d'acide nucléique (pyrimidines) et de composés analogues ont été mesurés à différentes valeurs de pH.

2. La variation des spectres avec le pH peut, dans tous les cas, être expliquée sur la base d'une dissociation ionique.

References p. 218.

3. Les constantes de dissociation apparente ont été calculées à partir des spectres.

4. La structure de l'uracile (et de la thymine) présente la forme dicéto en solution neutre. L'existence de deux equilibres ioniques en solution alcaline est démontrée et l'ordre des dissociations est respectivement des groupes 2- et 4-hydroxyles.

5. Les spectres et la structure d'autres dérivés pyrimidiques sont discutés.

ZUSAMMENFASSUNG

1. Die Ultravioletabsorptionsspektren einer Anzahl von Nukleinsäurederivaten (Pyrimidinen) und verwandten Verbindungen wurden bei verschiedenen pH-Werten gemessen.

2. Die Variation der Spektren mit dem pH kann in allen Fällen auf Grund einer Ionendissoziation erklärt werden.

3. Die scheinbaren Dissoziationskonstanten wurden von den Spektren abgeleitet.

4. Das Uracil (und Thymin) liegt in neutraler Lösung in der Diketo-Form vor. Zwei Gleichgewichtszustände wurden in alkalischer Lösung nachgewiesen; die Dissoziation geschieht erst an den 2- und dann an den 4-Hydroxylgruppen.

5. Die Spektren und die Struktur anderer Pyrimidinderivate werden diskutiert.

REFERENCES

1 W. STENSTROM AND N. GOLDSMITH, *J. Phys. Chem.*, 30 (1926) 1683.
2 P. A. LEVENE, L. W. BASS, AND H. S. SIMMS, *J. Biol. Chem.*, 70 (1926) 229.
3 J. R. MARSHALL AND J. WALKER, *J. Chem. Soc.*, (1951) 1004.
4 G. H. BEAVEN, E. R. HOLIDAY, AND E. M. JOPE, *Faraday Soc. Discussion*, No. 9 (1950) 406.
5 H. M. KALCKAR, *J. Biol. Chem.*, 167 (1947) 429, 445, 461.
6 G. E. HILBERT AND T. B. JOHNSON, *J. Am. Chem. Soc.*, 52 (1930) 2001.
7 T. B. JOHNSON AND I. MATSUO, *J. Am. Chem. Soc.*, 41 (1919) 782.
8 D. DAVIDSON AND O. BAUDISCH, *J. Am. Chem. Soc.*, 48 (1926) 2382.
9 G. E. HILBERT AND T. B. JOHNSON, *J. Am. Chem. Soc.*, 52 (1930) 1152.
10 G. E. HILBERT AND E. F. JANSEN, *J. Am. Chem. Soc.*, 57 (1935) 552.
11 T. B. JOHNSON AND R. D. COGHILL, *J. Am. Chem. Soc.*, 64 (1925) 233.
12 R. D. HOTCHKISS, *J. Biol. Chem.*, 175 (1948) 315.
13 G. R. WYATT, *Biochem. J.*, 48 (1951) 581.
14 P. A. LEVENE AND H. S. SIMMS, *J. Biol. Chem.*, 65 (1925) 519.
15 J. M. PLOESER AND H. S. LORING, *J. Biol. Chem.*, 178 (1949) 431.
16 J. E. AUSTIN, *J. Am. Chem. Soc.*, 56 (1934) 2141.
17 J. R. LOOFBOUROW, M. M. STIMSON, AND M. J. HART, *J. Am. Chem. Soc.*, 65 (1943) 148.
18 W. R. BRODE, *Chemical Spectroscopy*, John Wiley & Sons, Inc., N.Y., (1943).
19 A. M. MICHELSON, W. DRELL, AND H. K. MITCHELL, *Proc. Nat. Acad. Sci.*, 37 (1951) 396.
20 M. BACHSTETZ, *Ber.*, 63b (1930) 1000.
21 B. LYTHGOE, *Quart. Rev.*, 3 (1949) 181.
22 H. SPECKER AND H. GAWRASCH, *Ber.*, 75 (1942) 1338.
23 R. E. STUCKEY, *J. Pharm. Pharmacol.*, 1 (1949) 382.
24 W. C. SCHNEIDER AND I. F. HALVERSTADT, *J. Am. Chem. Soc.*, 70 (1948) 2626.
25 P. A. TRUBSBACH, *Z. Phys. Chem.*, 16 (1895) 725.
26 J. J. FOX AND D. SHUGAR, *Biochim. Biophys. Acta*, in press.
27 J. J. FOX AND D. SHUGAR, *Bull. soc. chim. Belges*, 61 (1952) 44.

Received December 7th, 1951