INFRARED SPECTRA AND TAUTOMERIC STRUCTURE OF NUCLEOSIDES AND NUCLEOTIDES IN D_2O SOLUTION. II

H. TODD MILES

Section on Enzymes, National Heart Institute, National Institutes of Health, Bethesda, Md. (U.S.A.)

The probable function of hydrogen bonding in the structure of nucleic acids[1] has increased the importance of obtaining definite evidence on the keto-enol tautomerism of the nucleosides and nucleotides.

Evidence was presented earlier[2] that uridine and, by analogy, thymidine, have the diketo structure in D_2O solution. The present paper employs infrared spectra of new model pyrimidine nucleosides[3] to confirm that thymidine does have the diketo structure in neutral solution and that cytidine has the amino rather than the imino structure.

A number of dihydropyrimidine nucleosides have been prepared by hydrogenation with a new rhodium catalyst[4] and studied spectrally. Possible biological importance of such compounds has been shown by the isolation of dihydrocytidylic acid from liver[5], and metabolic roles have been proposed by other workers[6], who have reported some of the same compounds. In addition to their usefulness for structure studies, it is possible that some of the other model compounds may be of direct biological interest. It has been found, for example, that 3-methyluridine inhibits the growth of *Neurospora* Strain 1298[7], and 6-dimethylamino-9-β-D-ribofuranosylpurine[8] was found to show some activity against transplanted adenocarcinoma of the C_3H mouse. 3-Methylthymidine, in which the labile hydrogen (necessary for hydrogen bonding in the DNA structure proposed by WATSON AND CRICK) is replaced by a methyl group, may prove interesting in that its lack of a hydrogen should make it incapable of incorporation into DNA in biological systems capable of incorporating thymidine.

EXPERIMENTAL*

Materials

The preparation of compounds I, II, III, 3-methylthymidine, and 1-β-D-glucopyranosyl-3-methyluracil (Table I) has been reported previously[3].

The following physical constants have been measured in neutral aqueous solution.

1-β-D-Glucopyranosyl-3-methyluracil, $[\alpha]_D^{24} = 18.0°$; λ max = 259 mμ; ε max = 9,380.

1-β-D-Glucopyranosyl-3-methylthymine $[\alpha]_D^{24} = 16.9°$; λ max = 264 mμ; ε max = 8,950.

3-Methylthymidine, $[\alpha]_D^{24} = 24.3$; λ max = 267 mμ; ε max = 9,100.

1-β-D-Glucopyranosyl-4-dimethylamino-2-pyrimidone (III), $[\alpha]_D^{24} = 12.5°$; λ max = 276 mμ; ε max = 13,800; λ max = 283 mμ in 0.1 *N* HCl; ε max = 15,500 in 0.1 *N* HCl.

Adenosine-5′-phosphate dihydrate was obtained from Pabst, and the monosodium salt was prepared by adding one equivalent of sodium hydroxide to a D_2O solution and lyophilizing the solution.

* Microanalyses by J. F. ALICINO, Metuchen, New Jersey. Melting points were measured on a Kofler apparatus.

References p. 52.

Deoxycytidine hydrochloride was obtained from Sigma, and deoxyuridine from the California Foundation for Biochemical Research.

The hydroperchlorate of III was prepared by adding an equivalent of 60% perchloric acid to an alcoholic solution of the nucleoside. The m.p. was 220–222°.

The spectrum of the hydroperchlorate (or deuteroperchlorate) of cytidine was obtained by making a 0.1 *N* solution of $HClO_4$ in D_2O and using this solvent for both sample and comparison cells. The small amount of water introduced with the 60% $HClO_4$ was not enough to interfere with the spectrum.

Hydrogenation

The rhodium-alumina catalyst reported by COHN AND DOHERTY[4] was used for hydrogenation of various pyrimidine nucleosides. The reductions were carried out in 10^{-3} *N* HCl in a semi-microhydrogenation apparatus at atmospheric pressure. In order to measure accurately the hydrogen uptake it was desirable to stir the catalyst (about one-fourth the weight of compound used) with hydrogen before adding the compound to be reduced through the side arm of the hydrogenation flask. At the completion of the reduction, the reaction mixture was filtered through celite or centrifuged to remove catalyst and then lyophilized (dry ice and NaOH traps in line). The material was crystallized, if possible, from a mixture of methanol with ethyl acetate or ether[3].

Dihydrouridine and dihydro-3-methyluridine were amorphous and deliquescent. Their infrared spectra are reported in Table I, but they were not further characterized.

Dihydroglucosyluracil had a m.p. of 243–244° and $[\alpha]_D^{24} = 9.7°$. Analysis: Calcd. for $C_{10}H_{16}N_2O_7$; C 43.48; H 5.84; N 10.14. Found: C 43.59; H 5.92; N 10.05.

Dihydro-3-methylglucosyuracil had a m.p. of 183–184° and $[\alpha]_D^{24} = 10.2°$. Calcd. for $C_{11}H_{18}N_2O_7$: C 45.51; H 6.25; N 9.65. Found: C 45.48; H 6.29; N 9.65.

Dihydrothymidine had a m.p. of 206–209° and, after three crystallizations, $[\alpha]_D^{24} = -25.6°$. Analysis: Calcd. for $C_{10}H_{16}N_2O_5$: C 49.17; H 6.60; N 11.47. Found: C 49.34; H 6.67; N 11.21. GREEN AND COHEN[6] have recently reported preparation of the same compound but found a m.p. of 152–153°. The two preparations may be polymorphous forms of the same substance or possibly a pair of diastereoisomers since a new asymmetric center is generated by the reduction. This second possibility might stem from asymmetric induction proceeding in different directions in neutral water solution and in 10^{-3} *N* HCl or from the isolation of different products from the

TABLE I

INFRARED ABSORPTION SPECTRA IN DEUTERIUM OXIDE SOLUTION

Compound	*Wave length in μ**
1-β-D-Glucopyranosyl-3-methyluracil	5.90 (S); 6.05 (VS); 6.19 (S)
Deoxyuridine	5.91 (S; Sh); 6.01 (VS)
3-Methylthymidine	5.93 (S); 6.00 (S); 6.18 (VS)
1-β-D-Glucopyranosylthymine	5.90 (VS); 6.00 (VS); 6.10 (VS)
1-β-D-Glucopyranosyl-3-methylthymine (I)	5.93 (S); 6.00 (S); 6.16 (VS)
1-β-D-Glucopyranosyl-4-ethoxy-5-methyl-2-pyrimidone (II)	6.00 (VS); 6.15 (S)
Cytidine	6.08 (VS); 6.21 (S)
Cytidine hydroperchlorate	5.85 (S); 6.01 (VS)
Deoxycytidine	6.05 (VS); 6.15 (S)
Deoxycytidine hydrochloride	5.83 (S); 6.01 (VS); 6.28 (W)
1-β-D-Glucopyranosyl-4-dimethylamino-2-pyrimidone (III)	6.09 (VS); 6.15 (S; Sh); 6.50 (VS)
1-β-D-Glucopyranosyl-4-dimethylamino-2-pyrimidone hydroperchlorate	5.84 (S); 6.04 (VS); 6.20 (M)
Adenosine-5′-monophosphate sodium salt	6.14 (VS); 6.31 (W)
6-Dimethylamino-9-β-D-ribofuranosylpurine (IV)	6.18 (VS)
Dihydrouridine	5.85 (S; Sh); 5.98 (VS)
Dihydro-3-methyluridine	5.85 (S); 6.02 (VS)
Dihydro-1-β-D-glucopyranosyluracil	5.82 (S); 5.87 (S); 5.98 (VS)
Dihydro-1-β-D-glucopyranosyl-3-methyluracil	5.80 (S); 5.99 (VS)
Dihydrothymidine	5.88 (S; Sh); 6.00 (VS)
Dihydro-1-β-D-glucopyranosylthymine	5.80 (M); 5.87 (S); 5.97 (VS)

* The following abbreviations are used in the table to give a rough indication of band intensities: VS, very strong; S, strong; M, moderate; W, weak. Sh indicates that the band appears as a shoulder on the side of another band.

References p. 52.

same reaction mixture by the use of different solvents (butanol in the former case and methanol ethyl acetate-ether in the latter).

Dihydroglucosylthymine had a m.p. of 138.5–141°. Despite recrystallization to constant m.p., the analysis indicates that the compound was not pure. Calcd. for C_{11} H_{18} N_2 O_7: C 45.51; H 6.25; N 9.65. Found: C 43.12; H 6.48; N 9.69.

The method of measuring the spectra in D_2O solution has been reported previously[2]. A Perkin-Elmer model 21 spectrophotometer was used.

RESULTS

The infared spectra of the new compounds are presented for the 5.5–6.5 μ range for the D_2O solutions in Fig. 1 and Table I.

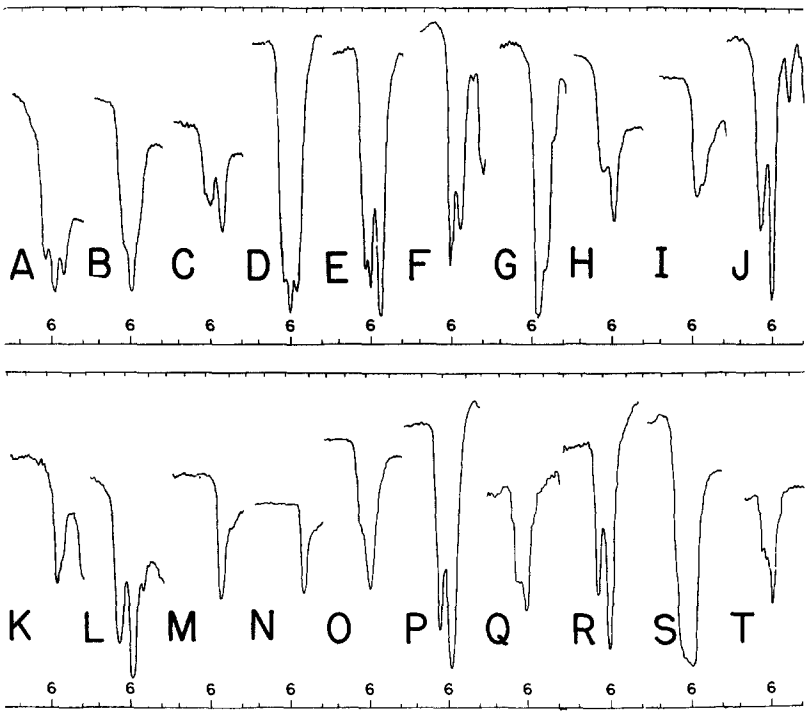


Fig. 1. Abscissa is wave length in mμ plotted against per cent transmission. Infrared spectra in D_2O of A, 1-β-D-glucopyranosyl-3-methyluracil; B, deoxyuridine; C, 3-methylthymidine; D, 1-β-D-glucopyranosylthymine; E, 1-β-D-glucopyranosyl-3-methylthymine (I); F, 1-β-D-glucopyranosyl-4-ethoxy-5-methyl-2-pyrimidone (II); G, cytidine; H, cytidine hydroperchlorate; I, deoxycytidine; J, deoxycytidine hydrochloride; K, 1-β-D-glucopyranosyl-4-dimethylamino-2-pyrimidone (III); L, 1-β-D-glucopyranosyl-4-dimethylamino-2-pyrimidone hydroperchlorate; M, adenosine-5'-monophosphate sodium salt; N, 6-dimethylamino-9-β-D-ribofuranosylpurine (IV); O, dihydrouridine; P, dihydro-3-methyluridine; Q, dihydro-1-β-D-glucopyranosyluracil; R, dihydro-1-β-D-glucopyranosyl-3-methyluracil; S, dihydrothymidine; T, dihydro-1-β-D-glucopyranosylthymine.

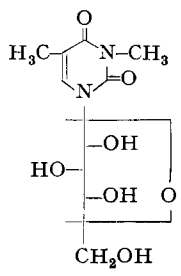
DISCUSSION

The evidence presented earlier[2] that uridine and its 5'-monophosphate have the diketo structure is considered conclusive since nontautomerizing derivatives of both members of the tautomeric pair were available. Uridine had a spectrum virtually

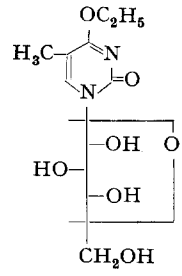
References p. 52.

identical with its 3-methyl derivative in the carbonyl region but different from that of a corresponding enol ether.

In the case of thymidine it was previously necessary to rely on analogy with uridine, but with the availability of (I) and (II) more direct evidence can be cited.



(I)

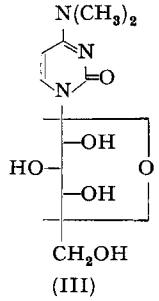


(II)

3-Methylthymidine was itself prepared[3] but was found to have too low a solubility in D_2O to give bands of satisfactory intensity at a cell thickness low enough to avoid D_2O bands. Use was, therefore, made of the glucose analogues shown above. In the present case, it can be seen that thymidine[2] and 1-β-D-glucopyranosylthymine have the same bands in the carbonyl region as do their respective N-methyl derivatives (Fig. 1; Table I).

A comparison of the spectra of 1-β-D-glucopyranosylthymine with those of (I) and (II) shows definitely that the diketo form predominates in solution; the bands of the former (5.90; 6.00; 6.10) are very similar to those of (I) (5.92; 6.00; 6.16) but different from those of (II) (6.00; 6.15), which lacks the shorter wave length carbonyl band. That the resolution of the thymidine solution spectrum is not as good as in the solid state, may indicate that a minor amount of the enolic form exists in solution, although the short wave length band of 3-methylthymidine in solution is also not well resolved. Since deoxyuridine (Table I) also shows a similar spectrum in solution, the poor resolution of the two bands appears to result from the effect of the deoxyribose on the pyrimidine ring. In any event, if any enolic form is present, the proportion is small, since the enolic form would have no shorter wave length band at all.

The availability of (III),



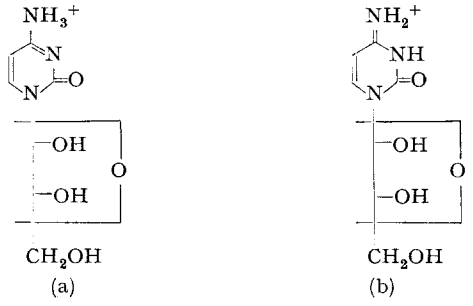
(III)

makes it possible to place on a stronger basis the previously expressed view[2] that cytidine and cytidylic acid sodium salt exist in the amino form in neutral solution. Cytidine has its carbonyl band at 6.08 μ and (III) has its band at 6.09 μ. While the lack of a corresponding model for the imino form leaves this demonstration less

References p. 52.

conclusive than the foregoing, it is, nevertheless, highly probable that the amino form is the only one present to an appreciable extent in the solution.

To study the structure of cytidine in acid solution, a crystalline perchlorate was prepared of (III) and the spectrum of cytidine was measured in $DClO_4$. Cytidine hydroperchlorate has bands at 5.85, 6.01 and 6.29, and the perchlorate of (III) has bands at 5.84, 6.04 and 6.20 (Table I).



It is, therefore, highly probable that cytidine and (III) have the same tautomeric form in acid solution, but, since (III) may also exist in acid in either of the above tautomeric forms, it is not possible to decide definitely between the two on the basis of the model alone. The results of the detailed X-ray analysis of adenine hydrochloride[9], however, may prove useful in interpreting the spectra. COCHRAN found that it was the N_1 that was protonated, and that N_{10} (the amino group external to the ring) has only two hydrogens attached. It is reasonable to believe, on the basis of analogy, that form (b) is the structure of cytidine in acid solution, and in fact the spectral bands can be interpreted much more readily on this basis than on that of form (a). Thus, the 6.07 μ band may be attributed to the $C=\overset{+}{N}$ grouping and the 5.88 μ band to the $C_2 = O$, shifted from its former position since the group is no longer in conjugation. If structure (a) were correct, the 6.07 μ band could be assigned to the carbonyl group, but it would be difficult to make a reasonable assignment of the 5.88 μ band.

Catalytic hydrogenations with rhodium catalyst[4] of a number of the compounds in this series were carried out and the spectra of the dihydropyrimidine nucleosides studied in D_2O solution (Fig. 1, Table I). The products were characterized spectrally, and, when crystalline, by optical rotation and elementary analysis (see Experimental).

A comparison of the spectra of the nucleosides (Table I and Ref. 2) and their dihydro derivatives shows that the band which appears at 6.1–6.2 μ in uridine and thymidine and their analogues has in every case disappeared in the dihydro compounds. Since this is in the region of C = C stretching vibrations[10], this band can very probably be assigned to the $C_5 = C_6$ double bond stretching vibration.

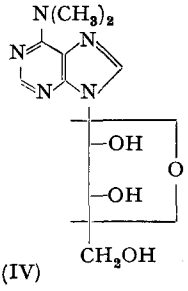
It appears further from a comparison of these spectra that with the diketo nucleosides there has been a slight shift to shorter wave length of the carbonyl bands. This shift is in the direction to be expected since the usual effect of conjugation is to shift a band to longer wave length, though in some cases the effect is small.

The question of the tautomeric structure of adenosine and adenylic acid in D_2O solution can be approached in the same manner. The necessary model compound 6-dimethylamino-9-β-D-ribofuranosylpurine (IV) was synthesized by KISSMAN,

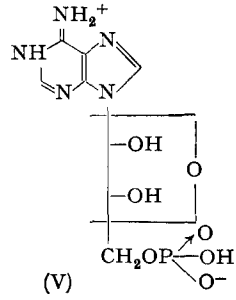
References p. 52.

PIDACKS AND BAKER[8]. (IV) has its principal peak in the double bond region at 6.18 μ in D_2O solution (Fig. 1 and Table I) and at 6.20 (VS), 6.35 (S) and 6.50 (M) in a mull.

Adenosine is too insoluble in D_2O to give a satisfactory spectrum, as is adenylic acid. The monosodium salt of adenylic acid, however, is soluble in D_2O and has its principal peak at 6.14 μ, indicating that sodium adenylate has the amino structure in D_2O and no charged ring nitrogen (a calculation from the pK values[11] in 0.1 N



(IV)



(V)

NaCl indicates a (RN <)/(RHN$^+$ <) ratio of about 25 in water). There is no firm basis at present for assigning these bands to particular C = N double bonds in the ring, but such an assignment is not necessary for an empirical correlation when adequate model compounds are available.

It is worth noting that the situation in the solid state appears to be different from that in solution. Adenylic acid has a very strong peak at 5.89 μ and no other peak of any prominence in the double bond region. The monosodium salt of this acid has peaks at 5.95 (VS), 6.01 (VS) (these two are not well resolved) 6.20 (VS), and 6.35 (M). The free acid presumably exists in the zwitterion form, which probably has the tautomeric structure (V), in view of the X-ray evidence for adenine hydrochloride[9] (the double bonds and the positive charge can, of course, occupy other positions in other forms contributing to the resonance hybrid). Adenosine itself has bands in a mull spectrum at 5.93 (VS), 6.20 (S), and 6.37 (W).

The mull spectrum of adenine has bands at 5.96 (S) and 6.19 (VS) and that of adenine sulfate at 5.83 (VS), 6.15 (M), and 6.33 (M). Thus, when the most prominent band in a mull spectrum of an adenine nucleoside or nucleotide appears at about 6.15, it very probably has the amino structure (on the basis of (IV), which can have no other structure). When the band appears at 5.82–5.89, it presumably indicates a positively charged ring of the structure shown above for adenylic acid (on the basis of the X-ray analysis of adenine hydrochloride). Those bands which fall between these extremes at 5.95–6.01 do not have so firm a basis for assignment, but their departure from the wave length observed for (IV) in either solid or solution and for sodium adenylate in solution is presumably due either to the presence of the imino structure in the solid state or, less probably, to the extensive hydrogen bonding which must be present in the solid.

ACKNOWLEDGEMENTS

The author is indebted to Miss PATRICIA WAGNER for measuring the ultraviolet spectra and specific rotations and to Mrs. KATHERINE WARREN for the mull spectra and some of the solution spectra.

References p. 52.

The gift of a sample of 6-dimethylamino-9-β-D-ribofuranosylpurine (IV) from Dr. HENRY KISSMAN of Lederle Laboratories is also gratefully acknowledged.

SUMMARY

Thymidine exists in the diketo form in deuterium oxide solution, and cytidine exists in the amino form. In acid solution, cytidine and cytidylic acid probably have an imonium structure. The monosodium salt of adenylic acid has the amino form in solution and, possibly, the imino form in the solid state. Adenylic acid has an imonium structure in the solid state.

REFERENCES

1 J. D. WATSON AND F. H. C. CRICK, *Nature*, 171 (1953) 737.
2 H. TODD MILES, *Biochim. Biophys. Acta*, 22 (1956) 247.
3 H. TODD MILES, *J. Am. Chem. Soc.*, 79 (1957) 2565.
4 W. E. COHN AND DAVID G. DOHERTY, *J. Am. Chem. Soc.*, 78 (1956) 2863.
5 L. GROSSMAN AND D. W. VISSER, *J. Biol. Chem.*, 216 (1955) 775.
6 M. GREEN AND S. S. COHEN, *J. Biol. Chem.*, 225 (1957) 397.
7 M. ROBERTS AND D. W. VISSER, *J. Biol. Chem.*, 194 (1952) 695.
8 H. M. KISSMAN, CHARLES PIDACKS AND B. R. BAKER, *J. Am. Chem. Soc.*, 77 (1955) 18.
9 W. COCHRAN, *Acta Cryst.*, 4 (1951) 81.
10 L. J. BELLAMY, *The Infra-red Spectra of Complex Molecules*, Methuen & Co., Ltd., London, 1954, p. 31.
11 *Pabst Laboratories Circular* OR-10, 1956 p. 20.

Received July 3rd, 1957

THE PROTEOLYTIC ENZYME SYSTEM OF SKIN

III. PURIFICATION OF PROTEINASE C AND ITS SEPARATION FROM AN INHIBITOR*

CHARLES J. MARTIN AND A. E. AXELROD

(with the technical assistance of ALBERT R. FRAZIER)

Biochemistry Department, University of Pittsburgh, School of Medicine, Pittsburgh, Pa. (U.S.A.)

The major portion of the proteolytic activity in high ionic strength extracts of rat skin acetone powder can undergo reversible inactivation upon exposure to an ionic strength of 0.6 to 0.8 for several hours[1]. The enzyme susceptible to such treatment was designated Proteinase C and its conversion from the inactive to the active state was readily attained by brief exposure to an ionic strength of 1.4 or greater. It was postulated[1] that the inactivation of Proteinase C occurred by association of the enzyme with an inhibitor, CIn, and that the reversal of inactivation was the result of Proteinase C–CIn

* This investigation was supported, in part, by a research grant (A-727) from the National Institute of Arthritis and Metabolic Diseases of the National Institutes of Health, Public Health Service and by the Office of Naval Research under Contract 1833 (00), NR 114-412.

References p. 62.