A Simple Method for the Preparation of ' Ribonucleoside Dialdehydes ' and Some Comments on their Structure

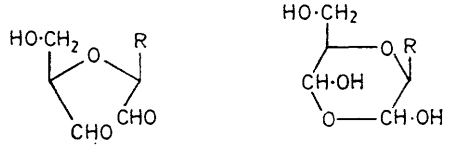
By **A. Stanley Jones,* Alex F. Markham,** and **Richard T. Walker,** Chemistry Department, The University of Birmingham, Birmingham B15 2TT

Adenosine, guanosine, inosine, cytidine, and uridine have been oxidised with periodate to give the respective α-substituted derivatives of α'-hydroxymethyloxydiacetaldehyde (' ribonucleoside dialdehydes '). These have been isolated by a simple procedure which involves extraction with ethanol, thus separating the dialdehydes from inorganic material. The compounds so isolated are more stable than has been suggested by early work in this field. I.r. and n.m.r. spectroscopy show that the compounds contain few if any free aldehyde groups. They appear to be mixtures of hydrated forms although they run as single spots on t.l.c. in several solvent systems. Dried samples also contain virtually no aldehyde groups, and the possibility that they are polymeric structures as suggested by others, is discussed. With the exception of ' guanosine dialdehyde,' the amplitude of the Cotton effect in the o.r.d. spectrum is less than that in the parent ribonucleoside.

OXIDATION of ribonucleosides with periodate gives α-substituted derivatives of α'-hydroxymethyloxydiacetaldehyde (I) (' ribonucleoside dialdehydes '). The initial interest in these compounds arose because of their importance in establishing the structure of the natural ribonucleosides.[1-3] Later it was shown that periodate oxidation of the terminal nucleoside of a ribonucleic acid followed by elimination of this residue under basic conditions provides a method for stepwise degradation of

[1] B. Lythgoe and A. R. Todd, *J. Chem. Soc.*, 1944, 592.
[2] J. Davoll, B. Lythgoe, and A. R. Todd, *J. Chem. Soc.*, 1946, 833.
[3] B. Lythgoe, H. Smith, and A. R. Todd, *J. Chem. Soc.*, 1947, 355.

a ribonucleic acid chain and hence a method of determining nucleotide sequence.[4,5] Terminal periodate oxidation of transfer ribonucleic acid has been used as a method of purification [6] and as a procedure for the attachment of fluorescent markers.[7] Ribonucleoside dialdehydes have also been used to prepare polynucleotide analogues.[8] In addition ribonucleoside dialdehydes show interesting biological properties, *e.g.* periodate-oxidised adenosine or adenosine 5′-phosphate inhibits nucleic acid synthesis in Ehrlich tumour cells, both *in vitro* and *in vivo*,[9] and periodate-oxidised adenosine 5′-triphosphate is a potent inhibitor of ribonucleotide reductase.[10] Inosine dialdehyde, which is undergoing clinical trials as an antitumour agent, shows activity against solid tumours.[11] In view of their importance, therefore, a simple method for the preparation of ribonucleoside dialdehydes in pure stable form is desirable.



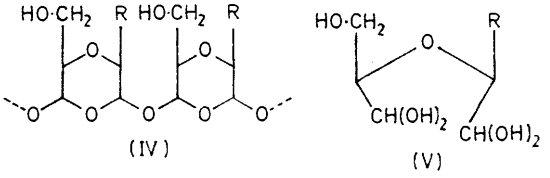
R = base residue

In the early work on these compounds they were usually isolated as derivatives.[2,3] This work gives the impression that the ribonucleoside dialdehydes are unstable compounds which are isolated with difficulty, but more recent work on adenosine dialdehyde indicates that this might not be so;[12] this compound can be recrystallised from water without decomposition.

In the present work we have confirmed that the ribonucleoside dialdehydes are easily isolated and stable. They were isolated by extraction from the dried reaction mixture with ethanol. The dialdehydes from adenosine, guanosine, and cytidine were finally obtained by crystallisation from water. The dialdehydes from inosine and uridine were too soluble in water for this to be done but they were obtained as pure freeze-dried solids. Although, from evidence quoted later, it appeared that the dialdehydes are mixtures, all attempts to separate these by t.l.c. failed; each ran as a single component in several solvent systems. Samples of these dialdehydes have been kept at room temperature for over 12 months and have shown no sign of decomposition. All are soluble in water to a concentration greater than 30 mg ml^{-1} and can be quantitatively reduced to the corresponding trialcohols by the literature procedure.[5] The reported instability of these dialdehydes could have been due to the introduction of traces of impurities from the ion-exchange resins used in their isolation.

The structures and reactivities of dialdehydes obtained by periodate oxidation of carbohydrates, including ribonucleosides, have been reviewed.[13] Addition of water across the two aldehyde groups to give the structure (II) has been proposed. Structure (III) has been put forward as a possibility for the ribonucleoside dialdehydes under acidic conditions [14] because under these conditions only one aldehyde group, that farther from the glycosidic centre, is reduced by sodium borohydride.

The most recent work on the ribonucleoside dialdehydes is that of Hansske *et al.*[12] on adenosine dialdehyde. These workers noted the greater hydrophobicity of the dialdehyde as compared with the parent nucleoside, and the absence of an aldehydic carbonyl absorption in the i.r. spectrum. These results indicated that some type of hydrated structure(s) was present. However, Hansske *et al.* also noted that by thorough drying the compound could be dehydrated completely, but that it still showed no aldehydic carbonyl band in the i.r. spectrum and no aldehydic proton signal in the n.m.r. spectrum. From this evidence and from certain features of the i.r. spectrum it was suggested that, in the anhydrous solid, the compound exists as the polymeric structure (IV) and that in concentrated solution this structure may persist and be in equilibrium with hydrated forms such as (II) and (V).



In general our results are similar to those of Hansske *et al.*[12] The chromatographic behaviour of all five ribonucleoside dialdehydes showed them to be more hydrophobic than the parent nucleosides, *i.e.* they had higher R_F values. The i.r. spectra showed no evidence

[4] (*a*) D. M. Brown, H. Fried, and A. R. Todd, *Chem. and Ind.*, 1953, 352; (*b*) P. R. Whitfield, *Biochem. J.*, 1954, **58**, 390.

[5] (*a*) K. Randerath, E. Randerath, L. S. Y. Chia, R. C. Gupta, and M. Sivarajan, *Nucleic Acids Res.*, 1974, **1**, 1121; (*b*) M. Sivarajan, R. C. Gupta, L. S. Y. Chia, E. Randerath, and K. Randerath, *ibid.*, p. 1329; (*c*) K. Randerath, L. S. Y. Chia, R. C. Gupta, E. Randerath, E. R. Hawkins, C. K. Brum, and S. H. Chang, *Biochem. Biophys. Res. Comm.*, 1975, **63**, 157.

[6] (*a*) J. S. Dixon and D. Lipkin, *Analyt. Chem.*, 1954, **26**, 1092; (*b*) H. von Portatius, P. Doty, and M. L. Stephenson, *J. Amer. Chem. Soc.*, 1961, **83**, 3351.

[7] S. A. Reines and C. R. Cantor, *Nucleic Acids Res.*, 1974, **1**, 767.

[8] M. G. Boulton, A. S. Jones, and R. T. Walker, *Biochim. Biophys. Acta*, 1971, **246**, 197.

[9] J. G. Cory, M. M. Mansell, C. B. George, and D. S. Wilkinson, *Arch. Biochem. Biophys.*, 1974, **160**, 495.

[10] J. G. Cory and C. B. George, *Biochem. Biophys. Res. Comm.*, 1973, **52**, 496.

[11] M. Slavik, *Ann. New York Acad. Sci.*, 1975, **255**, 266.

[12] F. Hansske, M. Sprinzl, and F. Cramer, *Bio-org. Chem.*, 1974, **3**, 367.

[13] R. D. Guthrie, *Adv. Carbohydrate Chem.*, 1961, **16**, 105.

[14] J. X. Khym and W. E. Cohn, *J. Amer. Chem. Soc.*, 1960, **82**, 6380.

for the presence of aldehydic carbonyl groups and the u.v. spectra were identical with those of the parent nucleosides. Our compounds were not so extensively dried as was the adenosine dialdehyde prepared by Hansske *et al.*, but they contained less than 1 mol. equiv. of water. The most detailed study was made on uridine dialdehyde. This gave a ^{1}H n.m.r. spectrum which was difficult to interpret and which indicated that the compound was a mixture of several structures. An aldehydic proton peak was present, however, but it corresponded to the presence of only about 5% of an aldehydic form. The ^{1}H n.m.r. spectrum of cytidine dialdehyde was too poorly resolved to be interpreted; no aldehydic proton peak was detected. The ^{13}C n.m.r. spectrum of uridine dialdehyde was also complex; it confirmed the results of the ^{1}H n.m.r. in that a peak which could be assigned to an aldehydic carbon atom was present, but was of low intensity.

Although the suggestion that the ribonucleoside dialdehydes exist in the polymeric form (IV) in the solid state and in concentrated solution appears reasonable, upon t.l.c. they all run as single spots with quite high R_F values. It appears, therefore, that in dilute, water-containing solutions the compounds are mainly mixtures of monomeric hydrated forms such as (II) and (V).

In o.r.d. spectra (Table 2) the amplitudes of the Cotton effects shown by the ribonucleosides are in good agreement with those reported,[15] as are the shapes of the curves. An artefact of the instrument used was that the zero cross-over points were slightly shifted relative to the published values. The amplitudes of the Cotton effects of the ribonucleoside dialdehydes were, with the exception of that of guanosine dialdehyde, smaller than those of the parent ribonucleosides. This could be attributed to less restriction of rotation of the base residues round the C–N glycosyl linkage than is the case in the ribonucleosides (see ref. 15), but may be merely a result of the presence of mixtures. Under the conditions used to measure the o.r.d. spectra it appears improbable that significant amounts of polymeric material are present.

TABLE 1

Properties of ribonucleoside dialdehydes

Dialdehyde from	R_F *	$\lambda_{max.}$/nm † (ε)	Formula	Found (%) ‡			Required (%)		
				C	H	N	C	H	N
Uridine	0.70	261 (10 100)	$C_9H_{10}N_2O_6,0.5H_2O$	42.8	4.4	11.4	43.1	4.4	11.2
Cytidine	0.45	271 (9 200)	$C_9H_{11}N_3O_5,0.5H_2O$	42.9	5.0	16.6	43.2	4.8	16.8
Adenosine	0.39	260 (14 900)	$C_{10}H_{11}N_5O_4,0.5H_2O$	43.9	4.4	26.3	43.8	4.4	25.5
Inosine	0.53	248 (12 200)	$C_{10}H_{10}N_4O_5,0.5H_2O$	43.2	4.0	20.2	43.6	4.0	20.4
Guanosine	0.55	253 (13 700)	$C_{10}H_{11}N_5O_5,2H_2O$	37.2	4.4	21.6	37.9	4.8	22.1

* In butan-1-ol–ethanol–water (4 : 1 : 5; organic phase). The compounds were also homogeneous on t.l.c. in propan-2-ol–ammonia (s. g. 0.88)–water (7 : 1 : 2) and in the two-dimensional cellulose t.l.c. system (E. Randarath, C. T. Yu, and K. Randerath, *Analyt. Biochem.*, 1972, **48**, 172). † Unicam SP 1800 instrument; solvent water; pH 5.0. ‡ After drying at 78 °C *in vacuo* (P_2O_5) for 8 h.

EXPERIMENTAL

Uridine Dialdehyde.—Uridine (2 g) was dissolved in water (100 ml), sodium periodate (1.8 g) in distilled water (80 ml) was added, and the solution was kept in the dark at 20 °C for 24 h. T.l.c. in butan-1-ol–ethanol–water (4 : 1 : 5; organic phase) then showed that complete oxidation had been achieved to give a single component with a higher R_F value than the original ribonucleoside (Table 1). The solution was evaporated to dryness under reduced pressure to give a solid which was twice extracted with ethanol [by treating the solid with boiling ethanol for 5 min; prolonged boiling with ethanol (15—30 min) gave a final product containing ethanol which could not be removed by drying]. Insoluble material was filtered off, the combined filtrates were evaporated to dryness, and the dry residue was again extracted with ethanol as above, a small amount of insoluble material being filtered off. The ethanolic solution was evaporated to dryness, traces of ethanol were removed by co-evaporation with water and the residue was dissolved in water (50 ml) to give a slightly cloudy solution; this was freeze dried to give uridine dialdehyde (2.02 g of moist material). After drying at 78 °C for 8 h *in vacuo* (P_2O_5) the compound gave the analytical results shown in Table 1. It ran as a single component on t.l.c. in butan-1-ol–ethanol–water (4 : 1 : 5; organic phase) (R_F 0.70) and in acetonitrile–water (21 : 4) (R_F 0.68).

The dialdehydes from inosine, adenosine, guanosine, and cytidine (Table 1) were obtained similarly. It was possible to obtain the dialdehydes from adenosine, guanosine, and cytidine in solid form by allowing them to separate from hot water upon cooling. The dialdehydes from uridine and inosine were too soluble for this. The solubilities of the dialdehydes of adenosine, guanosine, and cytidine in water at 18 °C were 36.7, 43.6, and 81.6 mg ml^{-1}, respectively.

I.r. Spectra.—These were measured (on samples dried at 78 °C for 3—4 h) as KBr discs with a Perkin-Elmer 180 spectrophotometer. All five ribonucleoside dialdehydes showed a broad band in the region 1 100—1 250 cm^{-1} (maximum 1 110 cm^{-1}) instead of the fine structure shown in this region by the ribonucleosides. No bands were found at 1 725, 2 740, or 2 855 cm^{-1} (for CHO). Bands in the region 1 650—1 695 cm^{-1} due to amide carbonyl groups of the bases were present except in the case of adenosine dialdehyde.

N.m.r. Spectra.—The ^{1}H spectra were measured on a Perkin-Elmer R14 (100 MHz) or Varian XL100 spectrometer and the ^{13}C spectrum on a Bruker AX-90E spectrometer (at the P.C.M.U., Harwell); $(CD_3)_2SO$ was used as solvent and tetramethylsilane as internal reference.

[15] T. L. V. Ulbricht in ' Synthetic Procedures in Nucleic Acid Chemistry,' vol. 2, eds. W. W. Zorbach and R. S. Tipson, Wiley-Interscience, New York, 1973, p. 177.

Uridine dialdehyde. The H-2′ and H-3′ resonances at δ 4.0 in the n.m.r. spectrum of uridine were not present in that of the dialdehyde. A doublet at δ 3.8 could be assigned to the 5′-protons and one at δ 5.7 to H-6. The spectrum in the region δ 5.0—5.5 was complex. The doublet at δ 7.87 in the spectrum of uridine (assigned to H-5) appears in the spectrum of the dialdehyde as a complex multiplet centred at δ 7.78. A small peak (*ca.* 0.05 H) at δ 9.72 could be ascribed to an aldehyde proton. A broad NH singlet occurred at δ 11.8. There were unassigned peaks at δ 5.88 and 6.11 (both <1 H). The ^{13}C spectrum was extremely complex in the region assignable to the carbon atoms of the base and the degraded sugar ring. There was however a peak of very low intensity at δ 201.46 p.p.m. which could be assigned to CHO.

Cytidine dialdehyde. A singlet at δ 7.16 was assigned to NH_2. The doublet at δ 7.83 in the spectrum of cytidine (H-5) appears in the dialdehyde spectrum as a complex multiplet, δ 7.5—7.9. There was no peak in the aldehyde region, and no other features were identified.

TABLE 2

O.r.d. spectra of ribonucleosides and ribonucleoside dialdehydes

Compound	$[\phi] \times 10^{-3}_{min.}$	$[\phi] \times 10^{-3}_{max.}$	Crossover λ/nm	Amplitude
Uridine	−4.7 at 250 nm	+6.4 at 280 nm	264	+11.1
Cytidine	−9.0 at 236 nm	+9.7 at 285 nm	267	+18.7
Adenosine	−1.6 at 275 nm	+3.2 at 245 nm	266	−4.8
Inosine	−3.6 at 260 nm	+0.7 at 220 nm	230	−4.3
Guanosine	−2.1 at 260 nm	+1.8 at 230 nm	250	−3.9
Urd dialdehyde	−8.1 at 255 nm	+0.8 at 280 nm	274	+8.9
Cyd dialdehyde	−2.7 at 230 nm	+5.7 at 287 nm	240	+8.4
Ado dialdehyde	−0.4 at 270 nm	+0.3 at 255 nm	266	−0.7
Ino dialdehyde	+2.2 at 260 nm	+3.2 at 245 nm		−1.0
Guo dialdehyde	−4.2 at 265 nm	+0.8 at 230 nm	243	−5.0

Because of their low solubility in dimethyl sulphoxide, the n.m.r. spectra of the other dialdehydes were not obtained. No attempt was made to record the spectra at high temperature (see ref. 12).

O.r.d. Spectra.—Spectra were recorded in water at pH 5 with an F.I.C.A. Spectropol 1 spectropolarimeter. The results are shown in Table 2.

We thank the S.R.C. for a studentship (to A. F. M.).

[6/176 *Received*, 27*th January*, 1976]