The Structure of Ribonucleic Acids

2. THE SMALLER PRODUCTS OF RIBONUCLEASE DIGESTION

By R. MARKHAM and J. D. SMITH
Agricultural Research Council, Plant Virus Research Unit, Molteno Institute, University of Cambridge

(*Received* 20 *February* 1952)

For many years it was thought that nucleic acids were tetranucleotides. This was largely due to observations such as those of Levene (1909) who showed that the ribonucleic acid (RNA) of yeast had approximately equivalent amounts of the two purines adenine and guanine and the pyrimidines cytosine and uracil. Analytical methods of that period were incapable of distinguishing small differences in the proportions of the bases, while the absence of techniques for isolating RNA in a native condition and for estimating the order of size of the molecules, was responsible for the failure of workers to realize that RNA was a much more complicated substance than could be represented by a chain of four nucleotides. Much effort was expended in testing this theory, for example by titration studies, made for the most part on degraded material and probably involving further degradation during the process. However, such experiments and those involving the action of enzymes led to the belief that RNA was a larger molecule having a 'statistical' rather than a 'structural' tetranucleotide composition (Gulland, Barker & Jordan, 1945). This hypothesis itself did not fit the observed facts, and the advent of chromatographic techniques has led to its complete abandonment. Reliable analyses for several ribonucleic acids are now available, and these fully confirm the earlier suspicions about the invalidity of the tetranucleotide hypothesis (Markham & Smith, 1950, 1951*a*; Smith & Markham, 1950).

Although much is known about the gross analytical chemistry of RNA much less is known about the nature of the internucleotide links. Cold alkaline hydrolysis leads to the formation of the '*a*' and '*b*' nucleotides (Carter & Cohn, 1950) which have been shown by Brown & Todd (1952) to be the nucleoside 2′- and 3′-phosphates (though not necessarily in that order). Cohn & Volkin (1951) have demonstrated the presence of nucleoside 5′-phosphates in digests after treatment of RNA with ribonuclease plus phosphodiesterase, thus showing that the 5′-hydroxyl group also participates in the internucleotide link. The quantity of 5′-phosphates isolated by these authors suggests that a large number, if not all, of the internucleotide linkages are through the sugar residues.

In the following paper we shall describe the smaller fragments produced by ribonuclease digestion—the mono-, di- and tri-nucleotides, their separation and their structure—and we hope to demonstrate that one kind of phosphate diester linkage is common to all.

Nomenclature

As we shall be dealing in this paper with a large variety of complex molecules, we have adopted a shorthand method of notation. As it is at present impossible to differentiate with certainty between the nucleoside 2′- and 3′-phosphates, we shall assume for convenience that any link involving one of these positions, is a 3′-hydroxyl derivative. Adenylic, guanylic, cytidylic and uridylic acids are represented by A, G, C and U respectively, while their (cyclic) monohydrogen phosphate derivatives (Markham & Smith, 1952) are represented by A!, G!, C! and U! respectively. In a dinucleotide the order in which the nucleotides are written is such that the second nucleotide has a free 3′-phosphate, or the equivalent, and is linked by its 5′-hydroxyl group to the preceding nucleotide. Thus GU! is the 2′:3′-cyclic uridylic acid 5′-ester of guanosine 3′-phosphate (the 5′-carbon being the only free position for the ester link in the cyclic uridylic acid), while GU is the similar dinucleotide containing uridine 3′-phosphate. UG, on the other hand, would represent the guanosine 3′-phosphate, 5′-ester of uridine 3′-phosphate, which, incidentally, is not found in ribonuclease digests.

MATERIALS AND METHODS

The nucleic acids used were prepared as in the preceding paper.

An impure but otherwise excellent turnip-yellow-mosaic-virus nucleic acid was also prepared by denaturing the virus in neutral solution with an equal volume of neutral ethanol, spinning off the protein and precipitating the nucleic acid by increasing the ethanol concentration to 75% (v/v). The preparation was then dried at room temperature. In this way one can be sure that no damage has been done to the acid- and alkali-labile groups in this nucleic acid.

Ribonuclease. Ribonuclease was obtained in a crystalline form from two sources and both preparations had very similar properties.

Phosphomonoesterase. Phosphomonoesterase was prepared from hypertrophic human prostate glands by grinding them with sand in 0·9% NaCl solution and allowing the suspension to autolyse under toluene for 18 hr. at 37°. The liquid was then freed from solids by centrifuging and the fraction of material precipitating between 0·5 and 0·75 saturation $(NH_4)_2SO_4$ was collected. This was further purified by salt precipitation followed by dialysis. The centrifuged solution was then stored in 0·1 M-acetate buffer, pH 5, with a few drops of $CHCl_3$ added to prevent bacterial growth. This enzyme dephosphorylated nucleotides and glycerophosphate rapidly, but had no detectable ribonuclease or diesterase activity.

Chromatographic solvent systems. For chromatographing we used four main solvent systems. (1) The *iso*propanol-ammonia system of the preceding paper (Markham & Smith, 1952), (2) sat. $(NH_4)_2SO_4$, 80 parts; *iso*propanol, 2 parts; 0·5 M-sodium acetate, 18 parts (v/v/v). (Markham & Smith, 1951*a*), (3) *n*-butanol saturated with water, 90 parts; formic acid, 10 parts (v/v) (Markham & Smith, 1949), and (4) *iso*propanol, 170 ml.; conc. HCl, 44 ml.; water to 250 ml. (Wyatt, 1951). Solvent 3 was used for detecting nucleosides, while solvent 4 was used for estimating the composition of the various fractions after hydrolysis in N-HCl for 1 hr. at 100° to give the purines and the pyrimidine nucleotides.

Ribonuclease hydrolysis. Hydrolysis was carried out as described in the previous paper. It will be appreciated that, as many of the substances isolated are themselves either rapidly or slowly digested by ribonuclease, one must be prepared to accept very small yields. On the other hand, if only the stable di- and tri-nucleotides are to be isolated, then the enzymic digestion may be allowed to go to completion. The dialysates from the digest were concentrated *in vacuo* and then chromatographed.

Analysis of fractions. The polynucleotides and mononucleotides when hydrolysed in N-HCl at 100° for 1 hr. give free purines and pyrimidine nucleotides which could then be identified and analysed quantitatively by chromatographing in solvent 4 (Markham & Smith, 1951*a*).

The internucleotide links were broken by treating the substances with N-NaOH or N-HCl overnight or by boiling in the presence of $BaCO_3$. The components could then be identified by running in solvents 2 and 4.

RESULTS

The initial chromatography

The digest (10–20 mg.) was put on to Whatman no. 3 paper in about 0·5–1 ml. of solution. This solution was applied in a line across the top of the paper about 6 cm. from the top edge, and some 25 cm. long. The paper was cut into notches at its lower end, and was then developed in solvent 1 until the band of the cyclic pyrimidine nucleotides nearly reached the bottom. (This may be observed quite easily by examining the chromatograms in ultraviolet light through a blue filter held before the eye, when the bands will appear dark (see Markham & Smith, 1951*a*, p. 402).) After drying off the solvent the chromatogram was printed on to photographic paper as previously described (Markham & Smith, 1949) but for a longer period (3–5 min.). A typical result is shown in Fig. 1. To obtain this result one may have to make minor adjustments to the solvent system, the ammonia concentration being the most important factor. The bands once located were cut out and eluted by allowing water to run along them. About 1 ml. of the extract was collected from each strip and this was then evaporated to dryness in a current of air at a temperature not exceeding 50°. This procedure fractionated the mixture into six main fractions numbered 1–6 in order of R_F values (Fig. 1). These fractions were not very sharp, but it must be remembered that there may be more than twenty individual substances on the paper.

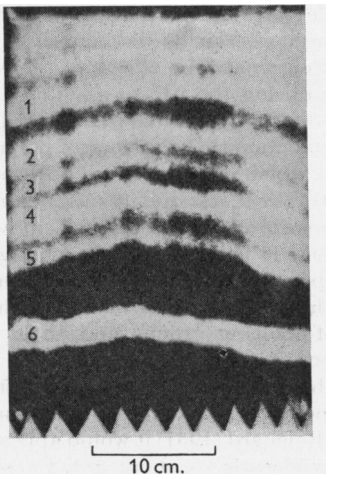


Fig. 1. The initial chromatogram. The bands are numbered as in the text. Band 4 is in the position to which yeast adenylic acid runs. Band 5*a* is not present in this particular chromatogram.

This chromatographic separation depends upon several factors. All guanine compounds move more slowly than do the similar compounds of the other bases, all cyclic derivatives have higher R_F values than the similar 2′- and 3′-phosphates, and mononucleotides move faster than dinucleotides, which in turn move faster than trinucleotides. This order is tabulated in Table 1.

Paper electrophoresis

The apparatus used has been described in detail in the first paper of this series, but it is necessary here to give a short account of certain theoretical points connected with its use.

The four nucleotides have each several ionizing groups, namely the primary and secondary OH groups of the phosphoric acid residues, the NH_2 groups of A, G and C, and the enolic OH groups of G, C and U. The last need not concern us here, but the first three are very important. The pK′ values (Levene & Simms, 1925) of the primary

phosphoric OH groups range from 0·7 to 1·02, so that for all practical purposes they can be regarded as fully dissociated at any pH value at which the internucleotide link is stable. The pK′ values for the NH_2 groups, on the other hand, are 3·7 for A, 2·3 for G and 4·24 for C, while U of course has none. Consequently, it is possible to select a pH value to take advantage of the net charges on the nucleotides to effect their separation. A useful pH value for this purpose is 3·5, though this is not suitable for all separations.

At pH values above 7, the secondary phosphoric OH groups are fully dissociated, so that compounds lacking these groups, i.e. polynucleotides containing cyclic nucleotides as end groups, or cyclic nucleotides themselves, can be distinguished. It is also possible to eliminate the effects of the NH_2 groups simply by raising the pH to a suitable value, i.e. above pH 5 for C and above pH 3·5 for G.

The electrophoretic behaviour of polynucleotides is a subject which we have already discussed briefly from a theoretical standpoint (Markham & Smith, 1951*b*). The rate at which a molecule moves through a fluid when it is exposed to a voltage gradient depends upon two main factors: (*a*) the net charge, which is the algebraic sum of the products of the number of ionizing groups and their percentage dissociation, and (*b*) the resistance to motion through the fluid. Nucleotides are rather asymmetrical molecules, and from the diffusion data of Gordon & Reichard (1951) it would appear that they have an axial ratio of about 5:1. Consequently a dinucleotide will probably have an axial ratio of between 2·5:1 and 10:1. Using these limits and assuming that the resistance to motion is the same as in translational diffusion, it is possible to set limits for the electrophoretic mobility of a nucleotide dimer as opposed to a monomer. The dimer has, of course, twice the charge per molecule, and rather less than twice the resistance to motion, so that the net effect is for any dimer to have a higher mobility than its monomer. We assess these limits as between 1·32 and 1·85 times that of the monomer, and for our preliminary work we took 1·54 as the most likely value. Since then we have been able to measure the actual values for the mono-, di- and trinucleotides of uracil and get a ratio of 1:1·27:1·37 for these substances; but these values contain an error in that the endosmotic flow is in the direction of motion, and consequently the absolute values of these ratios will actually be rather higher. A rough assessment of the amount of endosmotic flow suggests that about 25% of the movement of uridylic acid is due to this cause at pH 3·5, and the ratios of the actual mobilities thus become about 1:1·4:1·5. It will be seen that this series converges, so that a separation of tetra and higher nucleotides is unlikely.

In actual use, the electrophoretic separation has been simplified in that a preliminary fractionation has been achieved by chromatography, so that the system is not very complex. It is very useful, however, for the identification of various substances, to make use of the possibility of calculating relative mobilities, always bearing in mind that the method is to some extent empirical. As an example, take the relative net negative charges of the mononucleotides at pH 3·5. These are A, 0·46, G, 0·95, C, 0·16, and U, 1·00. From these one can calculate the mobilities relative to U of some di- and tri-nucleotides. Thus the mobility towards the positive electrode of

$$\text{AG} = (0{\cdot}46 + 0{\cdot}95)/1{\cdot}43 = 0{\cdot}99,$$
$$\text{of AC} = (0{\cdot}46 + 0{\cdot}16)/1{\cdot}43 = 0{\cdot}43,$$
$$\text{of AGC} = (0{\cdot}46 + 0{\cdot}95 + 0{\cdot}16)/2 = 0{\cdot}79,$$
$$\text{and of AAU} = (0{\cdot}46 + 0{\cdot}46 + 1{\cdot}0)/2 = 0{\cdot}96,$$

times that of uridylic acid at pH 3·5. The factors 1·43 and 2 represent the increased friction to motion of the dimers and trimers. (It is assumed that to a close approximation all the mononucleotides are about the same size.) With these figures it is possible to calculate the relative mobility of any mono-, di- and tri-nucleotide at this pH. For any other pH value, it is only necessary to calculate the charges for the four nucleotides from the appropriate pK′ values and proceed as above. It will be noticed that at this pH small inaccuracies in making up the buffer may make quite large differences to the mobility. In practice this is relatively unimportant.

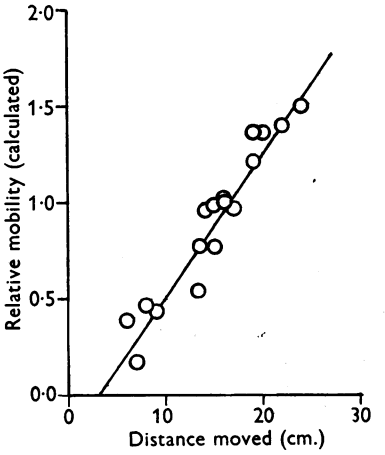


Fig. 2. Graph of mobility observed (in cm./2 hr./20 V./cm.) against relative mobility calculated from the charge and friction, for 19 mono- and poly-nucleotides (two of the points are on top of others).

Fig. 2 shows the extent to which one may rely on measurement of mobility for the identification of unknown nucleotides. This was compiled from a large number of different runs made for preparative

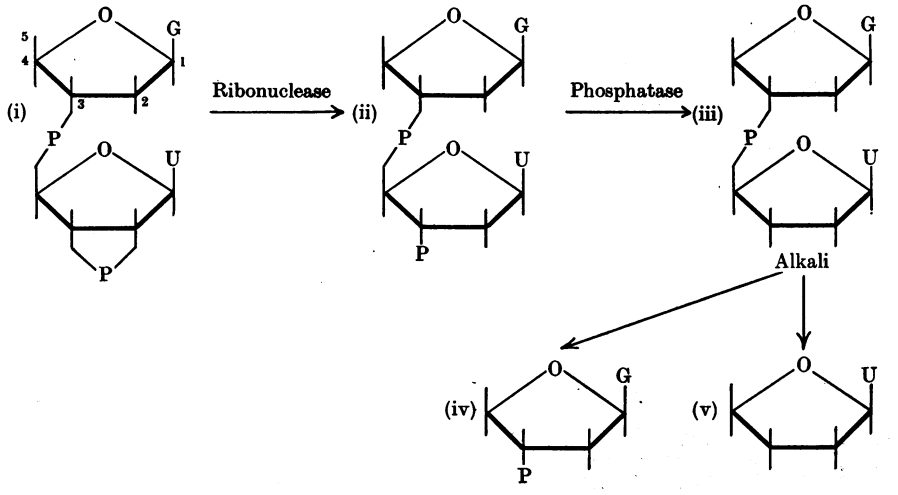
purposes, and no particular efforts were made to control the voltage gradient to better than ± 10 %.

For most purposes a run of 2 hr. with a gradient of 20 V./cm. at pH 3·5 is adequate. This gradient is given to a close approximation if a strip of filter paper 57 cm. (22·5 in.) long is used in the apparatus, the ends being allowed to dip into the electrode vessels about 3 cm. below the surface of the buffer and a total potential of 1000 V. is applied to the electrodes.

Cyclic dinucleotides

The isolation of the cyclic dinucleotides and their characterization is best illustrated by one example and for this we will take GU!. This substance was found in band 3 (Fig. 1) which, when dissolved in buffer and subjected to electrophoresis at pH 3·5 for 2 hr. at 20 V./cm. gave two bands moving about 20 and 15 cm. respectively. The former is GU!, the latter GC!.

band 3 was in fact GU! (i) which on ribonuclease treatment went to GU (ii). This was confirmed by running GU! before and after ribonuclease treatment in the electrophoresis apparatus at pH 7·4 (Fig. 3), when it was seen that digestion liberates an additional acidic group with a pK′ of between pH 4·5 and 6, i.e. a secondary phosphoric OH. This finding does not exclude the possibility of the secondary phosphoric OH being attached to the G residue, i.e. the substance being UG!, but this is unlikely as G! itself is not attacked by ribonuclease. The confirmation of the structure came from two different experiments. GU prepared from GU! by ribonuclease treatment gave on hydrolysis with barium carbonate a quantity of G!. GU on treatment with phosphomonoesterase gave rise to a new substance (iii) which had quite different properties (R_F value higher in solvent 1 and electrophoretic mobility much smaller) and which on hydrolysis in N-sodium hydroxide at room temperature gave rise



When analysed GU! was found to contain approximately equivalent amounts of guanine and uridylic acid (found: 0·9 mole/1·0 mole), and after alkaline hydrolysis was shown, by chromatography in solvent 2, to give guanylic acids '*a*' and '*b*' and a pyrimidine nucleotide. When treated further with ribonuclease, the dinucleotide gave a substance still containing equivalent amounts of guanine and uridylic acid (found: 0·9 mole/1·0 mole) but running in solvent 1 in the position of band 1. The R_F values in solvent 2 were reversed in order. This behaviour is similar to that of the cyclic pyrimidine nucleotides which, as we have shown in the previous paper, are digested by ribonuclease to liberate nucleoside 3′-phosphates. The rate of this digestion is much smaller than that of the internucleotide linkages and consequently it is possible to isolate these intermediates. So it was probable that the substance in

to equivalent quantities of guanylic acid (iv) and uridine (v). (iii) belongs to a new class of compounds, the dinucleoside monophosphates. There can therefore be no doubt that (i) represents the formula of GU! and (ii) that of GU.

Other dinucleotides

GC!, as we have already mentioned, was also found in band 3 (Fig. 1). On ribonuclease treatment it gave GC, which is to be found in band 1, and phosphatase subsequently converted it to the cytidine ester of guanylic acid. The analysis of GC gave G/C = 0·99 mole/1 mole. The separation of GC will be considered under trinucleotides.

AC and AU were found in chromatographic band 2 and moved about 8 and 16 cm. respectively in 2 hr. at 20 V./cm., on electrophoresis at pH 3·5 (Fig. 4). Analyses gave A/C = 1·03 moles/1 mole

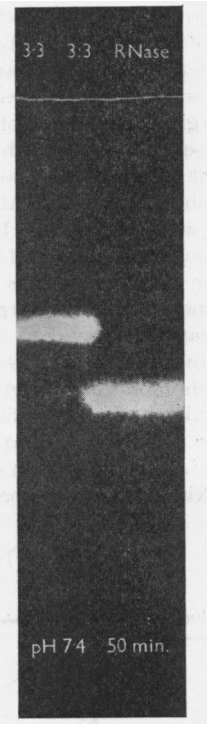


Fig. 3. Electrophoresis at pH 7·4 of GU! before and after ribonuclease treatment. It will be noted that the GU formed has an extra (secondary phosphoric OH) acidic group.

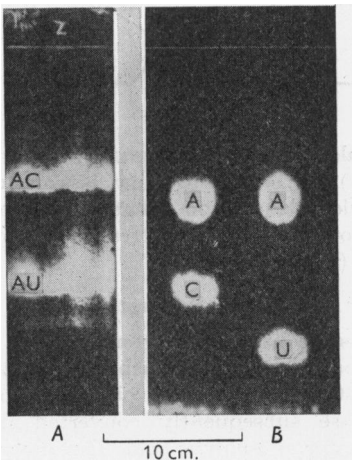


Fig. 4. Isolation and analysis of AC and AU. (*A*) The electrophoresis strip of band 2 (Fig. 1) showing AC and AU and a trace of GU from band 1; (*B*) the dinucleotides hydrolysed in N-HCl at 100° for 1 hr. and chromatographed in solvent 4, showing the presence of adenine in both and cytidylic acid in AC and uridylic acid in AU.

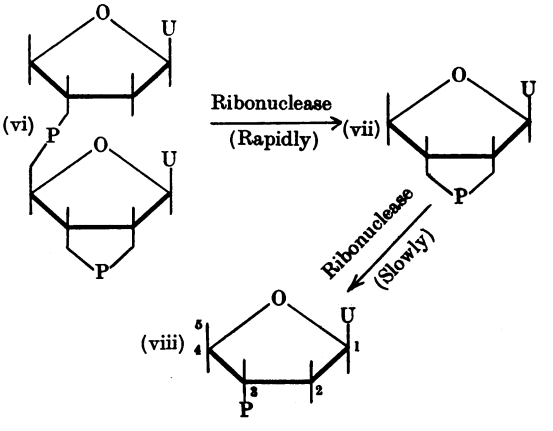
and A/U = 0·95 mole/1 mole. The analogous AC! and AU! were in chromatographic band 5.

The molar ratio of A/U! was found to be 0·97/1 and that of A/C! was 1·1/1. AC! and AU! were converted to AC and AU by ribonuclease, and the latter had only their pyrimidine nucleotide residues dephosphorylated by phosphatase, so that their structures are similar to that of GU. These dinucleoside phosphates have been isolated, hydrolysed in N-sodium hydroxide and degraded to purine nucleotides '*a*' and '*b*' and the appropriate pyrimidine nucleoside. The ratio of total purine nucleotide to pyrimidine nucleoside was found to be approximately 1/1 in all three cases.

AG (or GA) was found in band 1 (Fig. 1) in digests of commercial yeast nucleic acid. Its absence from carefully prepared specimens of nucleic acid suggests that it was an artifact of partial degradation of the nucleic acid, probably by alkali, which leaves this dinucleotide at one end of a short-chain fragment. The analysis gave a ratio of A/G = 1·1 moles/1 mole. It has not been subjected to phosphatase treatment and it may well be a mixture of AG and GA.

Diuridylic acid

Perhaps the most interesting dinucleotide is UU! This substance was found in chromatographic band 5 with the other cyclic non-guanine-containing dinucleotides, and its electrophoretic mobility at



pH 3·5 was approximately 1·4 times that of U and U!. UU itself was not found in the digests, as the internucleotide link breaks before the cyclic phosphate link. Consequently UU! (vi) was converted entirely to U! (vii) by ribonuclease before any U (viii) was liberated, and UU! was converted by both acid and alkali to U only, thus proving conclusively that the alkali-labile link between two pyrimidine nucleotides was the same as that between purine nucleotides and pyrimidine nucleotides (as in AU, etc.).

In addition to UU!, CC! was also found, particularly in digests of the turnip-yellow-mosaic-virus nucleic acid (which has a very high proportion of cytidylic acid), and we have also detected CU! (or UC!). These substances were found in chromatographic band 5, but little work has been done on them, as they are much more difficult to prepare than UU! and are only present in very small amounts.

Trinucleotides

The trinucleotides were found in chromatographic bands 1 and 2 with the exception of UUU!, which was in band 5.

UUU! has a mobility of 24 cm./2 hr. at pH 3·5 with a voltage gradient of 20 V./cm., and as may be expected was found in very low yield. Its identification is based upon the fact that it contains only uridylic acid, and must therefore be a polyuridylic acid. If it is assumed that each internucleotide link in such a substance is equally susceptible to ribonuclease, it will be seen that the probability of the higher polymers existing in any quantity decreases rapidly with chain length, so that the absence of a band between UU! and what we call UUU! makes it exceedingly unlikely that the latter has a more complex structure.

Table 1. *The chromatographic and electrophoretic properties of mono-, di- and tri-nucleotides*

Nucleotide	Chromatogram band no. (Fig. 1)	Mobility in cm./2 hr. at pH 3·5 and 20 V./cm.
A	4	8
G	2	14
C	4	6·5
U	4	16
A!	6	8
G!	(5*a*)	14
C!	6	7
U!	6	16
AC!	5	9
AU!	5	16
UU!	5	22
AC	2	8
AU	2	16
AG	1	15
GC!	3	15
GU!	3	20
GC	1	13·5
GU	1	19·5
UUU!	5	24
ACC!	2	6
AAC	1	13·5
AAU	1	17
AGU	1	19

Chromatographic band 1 subjected to electrophoresis at pH 3·5 for 2 hr. at 20 V./cm. gave rise to electrophoretic bands at (*a*) 13·5, (*b*) 17 and (*c*) 19 cm. (as well as one at 21 cm. which is GU). Band *a* was a mixture of GC and AAC. This could be resolved into its components by eluting and rerunning at pH 2·5 at 16 V./cm. for 3 hr., when GC moved 4·8 cm. and AAC moved 6·4 cm., or by running again at pH 5·0 for 3 hr. at 20 V./cm., when AAC ran just ahead (28 cm. against 24·5 cm.). The analysis of AAC gave a ratio of A/C = 1·92 moles/1 mole (theory 2/1). Band *b* was AAU (molar ratio A/U = 1·81/1). Band *c* was AGU (molar ratio A/G/U = 0·95/1/1·15). This trinucleotide could, of course, be GAU or a mixture of both AGU and GAU, but for reasons which we will discuss later it is probable that the uridylic acid residue is at one end of the molecule linked by its 5′-hydroxyl group.

Chromatographic band 2 subjected to electrophoresis at pH 3·5 for 2 hr. at 20 V./cm. sometimes gave a band moving 6 cm. This was ACC! (A/C = 1 mole/1·97 moles).

The positions and properties of these and some other substances are given in Table 1 which should prove an adequate guide to the isolation and identification of the various nucleotides.

DISCUSSION

The results presented in this paper and the previous one are all consistent with the hypothesis that those linkages in RNA which are broken by ribonuclease to give mono-, di- and tri-nucleotides, as well as the linkages in the latter which are broken by alkali, are all of the same type, i.e. a series of secondary phosphate ester links joining adjacent nucleosides through the hydroxyl groups at carbon atoms 3′ (or possibly 2′) and 5′. The link at C-5′ is evidently that which is broken by alkaline hydrolysis.

The end products of ribonuclease digestion have, with the exception of the mononucleotides and of AG, which is only obtained from chemically degraded RNA and is evidently an artifact, one and only one pyrimidine nucleoside residue per molecule, which is in the terminal position and is linked to the remainder of the molecule through its 5′-hydroxyl group, and is also esterified at C-3′ with phosphoric acid. Structures of this type are consistent with the observations of Schmidt, Cubiles & Thannhauser (1947) who reported that the action of phosphomonoesterase on the products of ribonuclease digestion is to hydrolyse off all the phosphoric acid stable to N-hydrochloric acid at 100° for 1 hr., i.e. that bound to pyrimidine nucleosides, while leaving all the rest of the phosphoric acid bound.

It will be seen that purine nucleotide-purine nucleotide linkages, and purine nucleotide 3′-pyrimidine nucleotide 5′-linkages are all stable to ribonuclease, as well as the 2:3′-cyclic phosphate linkages in cyclic purine nucleotides. The linkages which are broken are those which leave cyclic 2′:3′-

pyrimidine nucleotide residues at the end of a molecule, and this cyclic link is subsequently broken down itself. The existence of pyrimidine dinucleotides in partial digests and their subsequent hydrolysis by ribonuclease to give nucleoside 3′-phosphates through cyclic nucleotides, shows that it is unnecessary to postulate a branching chain to explain the formation of free pyrimidine nucleotides in the enzymic digests. These observations are consistent with the view that ribonuclease will only break secondary phosphate esters of pyrimidine nucleoside 3′-phosphates, that is to say, it is a highly specific phosphodiesterase which will only hydrolyse secondary phosphate esters of pyrimidine ribonucleoside 3′-phosphates. This specificity is reflected in its ability to hydrolyse the cyclic 2′:3′-monohydrogen phosphates of the pyrimidine ribonucleosides, which themselves have this structure.

Cohn (1951) has shown that all the pyrimidine nucleotides released from RNA by ribonuclease action are of the '*b*' form (which we have assumed to be the 3′-phosphates) and Brown, Dekker & Todd (1952) have shown that the end product of the ribonuclease digestion of the nucleoside 2′:3′-cyclic phosphates is also the '*b*' form, a finding which we have confirmed in the case of the cyclic cytidylic acid by running the digest in solvent 2, which separates the '*a*' and '*b*' cytidylic acids in a run of about 70 cm., the '*b*' form having the smaller R_F value. As a consequence of the fact that all the pyrimidine nucleotides found in those polynucleotides which are resistant to further ribonuclease digestion are at the ends of the molecules, it will be seen that they also will have been fixed by the enzyme action as the '*b*' isomers. Consequently, subsequent alkaline hydrolysis of such polynucleotides will only give the '*b*' pyrimidine nucleotides, although both forms of the purine nucleotides will be found.

As all the nucleoside residues in ribonuclease digests are substituted in the 3′ position, none of the substances described, excepting the dinucleoside monophosphates, react with periodate as do the nucleosides and their 5′-phosphate esters. Terminal nucleotide residues dephosphorylated with phosphomonoesterase will, however, react with periodate and this explains some of the observations made by Cavalieri, Kerr & Angelos (1951) and by Schmidt, *et al.* (1947) on RNA digested by both enzymes.

The large number of different polynucleotides obtained in digests of individual specimens of RNA leads one to the conclusion that either RNA is a very large molecule indeed, or that it consists of a large number of small molecules of varying composition. We are inclined to the latter view and in the next paper in this series will give evidence in support of this. In some of these small molecules a large number of pyrimidine nucleotide residues must be adjacent to each other, as a minimum of four is necessary in order to get a trinucleotide of the type of UUU! liberated from the middle of a chain as a result of ribonuclease digestion. It should be possible, if the rates of enzymic hydrolysis of the various internucleotide links were known, to calculate the relative numbers of adjacent pyrimidine nucleotide residues from the quantities of U, U!, UU! etc., and the similar cytidylic acid derivatives using probability theory. Whether this will be of any practical use is at the moment dubious, nor is there much point in devoting much time to a step-by-step degradation process of what is obviously a mixture. While these nucleic acids are clearly complex mixtures, we have evidence that they are not merely random mixtures and this evidence we will also present in the next paper.

It will be appreciated that, as so many polynucleotides exist in ribonuclease digests, the quantities of most of the individual polynucleotides are very small indeed. In fact most of the analyses reported in this paper have been made on a total of about 0·5 mg. of material, which in some cases has been subjected to two or more enzyme digestions and several chromatographic or electrophoretic separations. It is therefore not surprising that the accuracy of the analyses made on these substances is not quite as high as we have obtained on whole nucleic acid, but it will be noted that the accuracy is still of the same order as is obtained by other methods.

SUMMARY

1. The smaller products of ribonuclease digestion of ribonucleic acids have been isolated by a method involving a preliminary chromatographic separation followed by paper electrophoresis.

2. Detailed instructions for the isolation of fifteen of the smaller polynucleotides have been given.

3. The theory of the separation of polynucleotides by paper electrophoresis has been discussed.

4. Several dinucleotides have been subjected to a degradation process which shows that their general structure is: pyrimidine nucleoside 5′-(nucleoside-3′ or 2′-phosphate) ester 3′ or 2′-phosphate, or the cyclic form of such a dinucleotide: pyrimidine nucleoside 5′-(nucleoside-3′ or 2′-phosphate) ester 2′:3′-cyclic phosphate.

5. During ribonuclease digestion the cyclic forms of the dinucleotides are first liberated and the 2′:3′-cyclic phosphate link is subsequently slowly hydrolysed by the enzyme to give the 3′- (or 2′-) phosphates.

6. Trinucleotides liberated by ribonuclease action all have at least one pyrimidine nucleotide residue.

7. The alkali-labile link in ribonucleic acids is at C'-5.

8. Ribonuclease hydrolyses certain esters of pyrimidine ribonucleoside 3'- (or 2'-) phosphates.

9. Four dinucleoside monophosphates have been prepared by hydrolysis with ribonuclease followed by phosphomonoesterase.

10. A structure for ribonucleic acid has been proposed which satisfies these results on the basis of an unbranched chain.

REFERENCES

Brown, D. M., Dekker, C. A. & Todd, A. R. (1952). *J. chem. Soc.* p. 2715.

Brown, D. M. & Todd, A. R. (1952). *J. chem. Soc.* p. 44.

Carter, C. A. & Cohn, W. (1950). *J. Amer. chem. Soc.* **72**, 2604.

Cavalieri, L. F., Kerr, S. E. & Angelos, A. (1951). *J. Amer. chem. Soc.* **73**, 2567.

Cohn, W. E. (1951). *J. cell. comp. Physiol.* **38**, suppl. 1, 21.

Cohn, W. E. & Volkin, E. (1951). *Nature, Lond.*, **167**, 483.

Gordon, A. H. & Reichard, P. (1951). *Biochem. J.* **48**, 569.

Gulland, J. M., Barker, G. R. & Jordan, D. O. (1945). *Ann. Rev. Biochem.* **14**, 175.

Levene, P. A. (1909). *Biochem. Z.* **17**, 120.

Levene, P. A. & Simms, H. S. (1925). *J. biol. Chem.* **65**, 514.

Markham, R. & Smith, J. D. (1949). *Biochem. J.* **45**, 294.

Markham, R. & Smith, J. D. (1950). *Biochem. J.* **46**, 513.

Markham, R. & Smith, J. D. (1951*a*). *Biochem. J.* **49**, 401.

Markham, R. & Smith, J. D. (1951*b*). *Nature, Lond.*, **168**, 406.

Markham, R. & Smith, J. D. (1952). *Biochem. J.* **52**, 552.

Schmidt, G., Cubiles, R. & Thannhauser, S. J. (1947). *Cold Spr. Harb. Symp. quant. Biol.* **12**, 161.

Smith, J. D. & Markham, R. (1950). *Biochem. J.* **46**, 509.

Wyatt, G. R. (1951). *Biochem. J.* **48**, 584.

The Structure of Ribonucleic Acid

3. THE END GROUPS, THE GENERAL STRUCTURE AND THE NATURE OF THE 'CORE'

By R. MARKHAM and J. D. SMITH

Agricultural Research Council, Plant Virus Research Unit, Molteno Institute, University of Cambridge

(*Received* 20 *February* 1952)

Although much is known about the size and the various physical properties of the deoxyribonucleic acids, there is less information about the ribonucleic acids (RNA), largely because they are more difficult to prepare in an undegraded state, and to some extent also because they are easily and cheaply available commercially in a partially degraded form. In addition, RNA from many animal sources is inevitably subjected to enzyme action during isolation. The plant viruses offer what is perhaps the most suitable material for such investigations; some of them will yield nucleic acid under the mildest treatment (Markham & Smith, K. M., 1949). We also have some evidence that it is possible to prepare the RNA of yeast in a relatively undegraded form.

The determination of the molecular weights of such highly charged and asymmetrical molecules as those of the nucleic acids is a matter of extreme difficulty. Several attempts have been made to obtain this information for RNA (e.g. Tennent & Vilbrant, 1943; Bacher & Allen, 1950*a*), but it is likely that the results are only of the correct order of magnitude.

By more direct methods of a chemical nature, one can get an idea of the average molecular complexity. The various titration studies fall in this category (e.g. Fletcher, Gulland & Jordan, 1944), but these were made on partially degraded materials and probably involved some extra degradation as well. In this paper we shall describe how one can determine the average chain length of RNA and hence calculate the number of nucleotide residues per chain.

It has been known for some time that the course of digestion of RNA by ribonuclease stops before all the nucleotides are liberated. Kunitz (1940) found that 40% of his RNA was still precipitated by uranyl acetate after it had been digested to completion by the enzyme, while several workers have reported that a large fraction (about 10–25%) of the digested RNA is unable to dialyse through cellophan film (Bernheimer & Rodbart, 1948; Bacher & Allen, 1950*b*), although the observation of Greenstein, Carter & Chalkley (1947) to the contrary does not seem to have attracted much attention.

The 'non-dialysable' fraction of digested RNA has been called, among other names, the 'core' (e.g. Magasanik & Chargaff, 1951), and we propose to adopt this convention in this paper, although we hope to demonstrate that its non-diffusibility