Recognition and Estimation of 5-Methylcytosine in Nucleic Acids

By G. R. WYATT*
Agricultural Research Council Plant Virus Research Unit, Molteno Institute, University of Cambridge

(*Received* 19 *June* 1950)

The pyrimidine 5-methylcytosine was first reported as a constituent of a nucleic acid by Johnson & Coghill (1925), who prepared from hydrolysed tuberculinic acid a crystalline picrate which was identified as 5-methylcytosine picrate by microscopical examination. However, when Vischer, Zamenhof & Chargaff (1949) examined nucleic acid from avian tubercle bacilli by chromatography, they could not detect this substance. Nevertheless, Hotchkiss (1948) noted on paper chromatograms of hydrolysed thymus nucleic acid a small spot whose ultraviolet spectrum and chromatographic behaviour led him to suggest that it might be 5-methylcytosine.

As already reported in a preliminary note (Wyatt, 1950), I have found by a simple chromatographic method that 5-methylcytosine seems to occur in constant amounts in deoxypentose nucleic acids (DNA) from animals and from at least one higher plant, but has not so far been found in DNA from microbial sources. The isolation, partial characterization and estimation of this substance will now be described. Quantitative results will be presented along with the method and results of estimating the other pyrimidine components of some DNA's in the succeeding paper (Wyatt, 1951).

EXPERIMENTAL

Isolation of the new substance

The purines and pyrimidines set free by hydrolysis of DNA with formic acid (175° for 30 min.) were separated on paper chromatograms run with aqueous 65% (v/v) *iso*propanol 2N in respect of HCl as described by Wyatt (1951), and their positions detected by the photographic technique of Markham & Smith (1949). In preliminary experiments with DNA from ox spleen and from herring sperm, it was noted that in addition to guanine, adenine, cytosine, thymine and sometimes a trace of uracil, another substance formed a faint spot which moved slightly more rapidly on the chromatograms than cytosine. Tests had shown that the substance was not cytosine deoxyriboside, since its concentration was not appreciably diminished by increasing the hydrolysis time, and was not an aromatic amino-acid, since it gave no colour with ninhydrin. The isolation of a sample was undertaken as follows.

40 g. of crude herring-sperm DNA were freed of purines by methanolysis with HCl (Levene & Bass, 1931, p. 110). The filtrate from the precipitated purine hydrochlorides was evaporated to dryness, dissolved in 6N-HCl, and heated at 125° for 2 hr. to complete liberation of the pyrimidines. The hydrolysate was clarified with activated charcoal, evaporated to dryness under diminished pressure, redissolved in water and again evaporated. The tarry residue was taken up in water and freed of H_3PO_4 by neutralization with $Ba(OH)_2$ and filtration. After evaporation to a small volume, a crystalline precipitate separated out, consisting mainly of thymine; after further reduction in volume a precipitate containing cytosine, thymine and some of the new substance formed. The filtrate was made alkaline with NaOH and the remaining pyrimidines were precipitated with $AgNO_3$. After removal of Ag^+ with HCl this fraction was found to contain the new substance along with some thymine, uracil, adenine and much cytosine.

This concentrate was fractionated by paper chromatography. The solution was applied as a band across the top of sheets of Whatman no. 3 filter paper, which were run as chromatograms in *iso*propanol-aqueous HCl. In order to shorten exposure time, because of the thickness of Whatman no. 3 paper, Ilford Document Paper, no. 60, was used for printing these chromatograms instead of Reflex Document Paper, no. 50. As the chromatograms were overloaded (about 10 mg. of each substance can be separated on an 18 × 22 in. sheet of Whatman no. 3 paper), resolution was poor, but the bands containing the new substance were eluted and the eluate chromatographed in *n*-butanol-aqueous NH_3. The eluate this time contained only one ultraviolet-absorbing substance, and on evaporation deposited microscopic crystals.

Comparison with 5-methylcytosine

On paper chromatograms the new substance moved slightly more rapidly than cytosine in all the solvents tested, which included *iso*propanol-aqueous HCl, *n*-butanol-water, *n*-butanol-aqueous formic acid or NH_3, and *iso*amyl alcohol-aqueous NH_3. The position relative to cytosine was similar regardless of pH, but the separation from cytosine increased the higher the alcohol used. This strongly suggested a compound bearing ionizable groupings identical with those of cytosine, but differing by an additional alkyl group, and 5-methylcytosine seemed a likely possibility.

An authentic sample of 5-methylcytosine was not at the time available, but some synthetic 2-amino-4-hydroxy-5-methylpyrimidine (5-methyl*iso*cytosine) was kindly given by Prof. A. R. Todd, F.R.S. This had absorption spectra different from those of the unknown, and on chromatograms moved slightly faster.

If the unknown were 5-methylcytosine, it would be converted to thymine by deamination with HNO_2. Accordingly, 0·5 mg. of the substance in 1 ml. water was mixed with 2M-$Ba(NO_2)_2$ (2 ml.) and glacial acetic acid (0·5 ml.). After 6 hr. at room temperature, the excess acetic acid was removed by evaporation, and Ba^{++} with H_2SO_4. The filtrate from $BaSO_4$ was evaporated to dryness and dissolved in 0·1 ml.

* Present address: Laboratory of Insect Pathology, Box 490, Sault Ste Marie, Ontario, Canada.

of water, and spots were applied to paper for chromatography. A portion of the synthetic 5-methyl*iso*cytosine was treated in the same way. Both were found to have been converted quantitatively to a substance identical with thymine in its chromatographic movement (Fig. 1) and ultraviolet absorption spectrum.

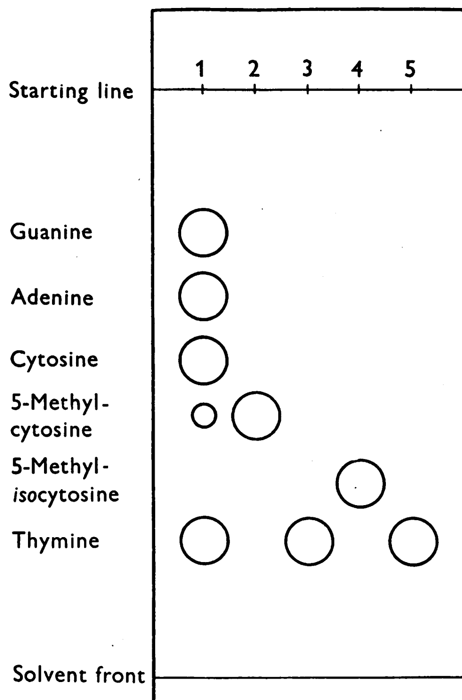

Fig. 1. Diagram of a paper chromatogram run in *iso*-propanol-aqueous 2 N-HCl. Solutions applied at positions 1–5 were: (1) hydrolysate of herring-sperm deoxypentose nucleic acid (DNA); (2) supposed 5-methylcytosine isolated from DNA; (3) product of deamination of (2); (4) synthetic 5-methyl*iso*cytosine; (5) product of deamination of (4).

There are only three substances which would yield thymine on deamination: 5-methylcytosine, 5-methyl*iso*-cytosine and 2:4-diamino-5-methylpyrimidine. The unknown had been found different from the second of these, and was most unlikely to be the third, since the possession of an additional amino group would cause slower movement in acid and more rapid movement in ammoniacal solvents on the chromatograms. This made it reasonably certain that the new substance was 5-methylcytosine.

After these experiments had been completed, a sample of 5-methylcytosine synthesized by the method of Wheeler & Johnson (1904) was obtained through the kindness of Dr A. Pircio, Fordham University, New York. This proved to be identical with the natural substance in its movement in both acid and ammoniacal chromatographic solvents and in its ultraviolet absorption spectra in neutral, acid and alkaline solution (Fig. 2). These spectra are also identical with those recently published by Hitchings, Elion, Falco & Russell (1949) for 5-methylcytosine prepared by a new synthesis.

Since the absorption curves and chromatographic behaviour provide evidence for the identity of the natural and synthetic substances, isolation of a sample sufficient for accurate elementary analysis has not been attempted. Owing to the low capacity of paper chromatograms, only a small portion of the concentrate obtained by precipitation had been fractionated by chromatography. Separation on starch columns was also tried, but their capacity proved

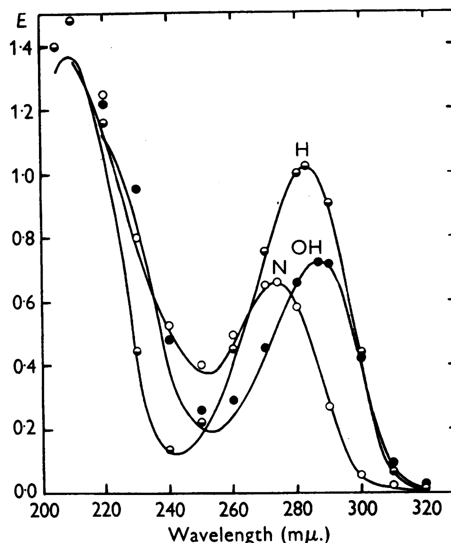

Fig. 2. Ultraviolet absorption spectra of 5-methylcytosine in 0·1 N-HCl (◒), in 0·1 N-NaOH (●) and at pH 7·0 (○). The curves represent the absorption of synthetic 5-methylcytosine (0·013 mg./ml.), and the points that of natural 5-methylcytosine eluted from chromatogram spots and read against paper blanks.

little greater than that of filter paper. Fractional crystallization proved useless for separation of cytosine and 5-methylcytosine. If it were desired to isolate a greater amount of natural 5-methylcytosine, a better starting material would be DNA from wheat germ, which contains one-third as much 5-methylcytosine as cytosine. The final fractionation might be accomplished by counter-current distribution or on ion-exchange columns.

Ultraviolet absorption characteristics of 5-methylcytosine

For determination of its ultraviolet absorption characteristics, some synthetic 5-methylcytosine was purified through the picrate, recrystallized from methanol and from water, and dried at 110° for weighing. Standard solutions were made up in water (taken to pH 7·0 with a little NaOH), 0·1 N-HCl and 0·1 N-NaOH. Molecular extinction coefficients at maxima and minima are given in Table 1. In acid solution there is an additional maximum at about 211 mμ. The characteristic shifts in acid and alkali correspond closely to those of cytosine, the wavelengths of absorption maxima being in each case longer by 6–8 mμ., and those of the minima by 3–4 mμ. The shift from the 'acid' to the 'neutral' spectrum takes place at a pH just below 7, and that from

the 'neutral' to the 'alkaline' at about pH 12, so that the curve for pH 7 (Fig. 2) and that for pH 11 (Hitchings *et al.* 1949) are almost identical.

Table 1. *Ultraviolet absorption characteristics of 5-methylcytosine*

(ϵ is molecular extinction coefficient.)

	Maxima		Minima	
Solvent	Wavelength (mμ.)	ϵ	Wavelength (mμ.)	ϵ
Water	274	6260	252	3680
0·1 N-HCl	283	9810	242	950
0·1 N-NaOH	287	6870	253	1670

Method of estimation

The amount of methylcytosine in animal nucleic acids is so small that estimation by elution of the spot on chromatograms of unfractionated hydrolysate is inaccurate, especially if traces of absorbing impurities are present. A procedure has therefore been used which is in effect two-dimensional chromatography, the substances being concentrated after running in the first dimension. Hydrolysate of 6–8 mg. of DNA is applied in a band 25 cm. long across the top of a sheet of paper, which is then run in *iso*propanol-aqueous-HCl, dried and printed. A transverse band containing all of both cytosine and methylcytosine is cut from the chromatogram; one end of this band is cut to a point, and the other end is placed in a trough containing water in a chromatogram tank. The pyrimidines move with the water front and are eluted quantitatively in the first 0·5 ml. to drip from the bottom of the paper strip, which is collected in a small tube. This is then evaporated to dryness and redissolved in 0·1 N-HCl (0·04 ml.). Measured volumes are placed as spots on filter paper and run in *n*-butanol-aqueous NH_3, in which cytosine and methylcytosine are well separated from each other and from impurities extracted from the paper of the first chromatogram. The separated spots are eluted and their pyrimidine contents estimated on the Beckman spectrophotometer; the amount of cytosine serves as a standard for referring the amount of methylcytosine to the whole nucleic acid. In this way methylcytosine can be estimated when it comprises as little as 0·1 % of the original nucleic acid.

Distribution of 5-methylcytosine

The various animal DNA's so far analysed (including mammals, a fish, an insect and an echinoderm) all contain methylcytosine, in amounts characteristic of the species source, and varying from 0·008 to 0·075 mol./4 mol. of nucleotide. The only DNA from a higher plant analysed, that of wheat germ, contains much more, 0·23 mol./4 mol. of nucleotide (Wyatt, 1951). None could be detected in DNA from the following microbial sources: human and bovine tubercle bacilli, *Escherichia coli*, *Esch. coli* bacteriophages T_2 and T_5, and the virus of the polyhedral disease of gipsy moth larvae; none or extremely little is present in DNA from baker's yeast. Nor could any be found in the ribonucleic acids of yeast and of turnip yellow-mosaic virus.

DISCUSSION

In the present state of knowledge as to the structure and function of nucleic acids nothing can be said as to the possible function of 5-methylcytosine. The amounts in which it occurs, however, varying with the source but constant from a given source, suggest that it is an essential constituent of certain DNA's, and no accident of enzyme action. The demonstration that these nucleic acids do contain this pyrimidine in addition to their well known components extends the evidence of recently published analyses that their structure is a matter of considerable complexity and worthy of much further investigation. In particular, it would be interesting to know whether the observation that methylcytosine is present in higher organisms but lacking in micro-organisms holds generally, and if so, at what point in the evolutionary scale it first appears.

It is ironical that the nucleic acid of tubercle bacilli, from which Johnson & Coghill (1925) originally reported 5-methylcytosine, should be one of those which do not contain it. As a clue to the source of their error, it is interesting to note that Vischer *et al.* (1949) examined by chromatography a specimen of methylcytosine obtained from Prof. T. B. Johnson and found it to contain 42 % of cytosine, so that the crystalline '5-methylcytosine picrate' with which the picrate from tuberculinic acid was compared may well have been chiefly that of cytosine.

SUMMARY

1. An aminopyrimidine which on deamination gives rise to thymine has been discovered as a new constituent of some deoxypentose nucleic acids (DNA). The ultraviolet absorption spectra and chromatographic behaviour of this substance are identical with those of synthetic 5-methylcytosine, and the tentative conclusion is drawn that it is the latter.

2. By a simple chromatographic technique, this pyrimidine can be estimated when occurring in amounts as small as 0·1 % of a nucleic acid.

3. Ultraviolet absorption spectra of 5-methylcytosine have been determined.

4. The new pyrimidine occurs in amounts characteristic of the source in all of the animal DNA's and the one plant DNA analysed, but is apparently lacking from DNA of bacteria and viruses, and from ribonucleic acids.

I wish to thank Dr J. D. Smith for information on his analyses of nucleic acid from *Esch. coli* and its bacteriophages, Dr R. Markham for helpful advice and discussion, and Dr K. M. Smith, F.R.S., for providing facilities for the work. The work was carried out while I was on transfer of research from the Forest Insect Investigations Unit, Division of Entomology, Department of Agriculture, Canada, to the Agricultural Research Council Plant Virus Research Unit, Molteno Institute, Cambridge. This paper is Contribution No. 2701, Division of Entomology, Science Service, Department of Agriculture, Ottawa, Canada.

REFERENCES

Hitchings, G. H., Elion, G. B., Falco, E. A. & Russell, P. B. (1949). *J. biol. Chem.* **177**, 357.
Hotchkiss, R. D. (1948). *J. biol. Chem.* **175**, 315.
Johnson, T. B. & Coghill, R. D. (1925). *J. Amer. chem. Soc.* **47**, 2838.
Levene, P. A. & Bass, L. W. (1931). *Nucleic Acids.* New York: Chemical Catalog Co., Inc.
Markham, R. & Smith, J. D. (1949). *Biochem. J.* **45**, 294.
Vischer, E., Zamenhof, S. & Chargaff, E. (1949). *J. biol. Chem.* **177**, 429.
Wheeler, H. L. & Johnson, T. B. (1904). *Amer. chem. J.* **31**, 591.
Wyatt, G. R. (1950). *Nature, Lond.*, **166**, 237.
Wyatt, G. R. (1951). *Biochem. J.* **48**, 584.

The Purine and Pyrimidine Composition of Deoxypentose Nucleic Acids

By G. R. WYATT*

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

(*Received* 19 *June* 1950)

Although the composition of nucleic acids has been subject to frequent investigation since these compounds were first isolated by Miescher in the last century, it is only through the application of techniques developed in the past few years that their accurate quantitative analysis has become possible. Early analyses by macrochemical methods (Steudel, 1906; Levene & Mandel, 1908) indicated that nucleic acids contained two purine and two pyrimidine bases in approximately equimolecular proportions. This gave rise to the tetranucleotide hypothesis, upheld by Levene & Bass (1931) and widely accepted, according to which the unit of nucleic acid structure is a group of four nucleotides comprising one of each of the four bases. Recent micro-analyses, however, by such methods as filter-paper chromatography and ultraviolet spectrophotometry, have shown that the composition of nucleic acids is not consistent with any such simplified theory, a conclusion which is in line with biological evidence of their intimate function in cell physiology and of the specific activities of certain of them, e.g. in inducing mutations in bacteria (Avery, MacLeod & McCarty, 1944; Boivin, 1947).

The feasibility of separating nucleic acid derivatives by paper chromatography was demonstrated by Vischer & Chargaff (1948*a*) and by Hotchkiss (1948). A method for the quantitative analysis of ribonucleic acids (RNA) was developed by Vischer & Chargaff (1948*b*), and later applied to deoxypentose nucleic acids (DNA) (Chargaff, Vischer, Doniger, Green & Misani, 1949). This involves separation of the purines and the pyrimidines: the former are precipitated on methanolysis with dry hydrogen chloride, and liberation of the latter is completed by formic acid at 175°. The two groups are then estimated on separate chromatograms.

A simpler procedure for the analysis of RNA has been described by Smith & Markham (1950), who use a short hydrolysis in N-hydrochloric acid to liberate purines along with pyrimidine nucleotides, all of which are then separated on one-dimensional paper chromatograms. This method is not applicable to DNA, from which nucleotides or nucleosides cannot be obtained quantitatively by chemical hydrolysis because of the lability of the deoxy-sugar. In the method now described, DNA is hydrolysed by formic acid to purines and pyrimidines, which are separated on single paper strips. Preliminary results from this method have already been reported (Wyatt, 1950), and the identification and estimation of the pyrimidine 5-methylcytosine in DNA is described in the preceding paper (Wyatt, 1951).

Since this work was done, Daly, Allfrey & Mirsky (1950) have published analyses of DNA from a number of sources, based on separations obtained with starch columns. Their results differ from those now presented in that they failed to recognize 5-methylcytosine, and inclined to the view that the differences between various nucleic acids were not significant.

METHODS

Detection of the spots of chromatograms

For detection of the spots of purines and pyrimidines on paper chromatograms the photographic technique of Markham & Smith (1949) was used, which consists in making contact prints of the dried chromatograms on reflex document paper with filtered ultraviolet light of wavelengths 254 and 265 mμ.

Chromatographic solvents

Although purines and pyrimidines can be separated by various neutral, ammoniacal or weakly acid solvents (a number are listed by Markham & Smith, 1949), such solvents

* Present address: Laboratory of Insect Pathology, Box 490, Sault Ste Marie, Ontario, Canada.