479. *The Nucleotide Sequence in Deoxypentosenucleic Acids. Part IV.* The Deoxyribonucleic Acid of* Mycobacterium phlei.

By A. S. Jones, M. Stacey, and Barbara E. Watson.

The deoxyribonucleic acid of *Mycobacterium phlei* has been isolated, characterised, and treated with mercaptoacetic acid at 37° in the presence of zinc chloride and anhydrous sodium sulphate. The resulting *aldehydo*-apurinic acid di(carboxymethyl) dithioacetal was degraded at pH 12 at 37° and the components so produced were separated into 20 fractions by paper chromatography and paper electrophoresis. Eight components have been identified and their proportions in the total hydrolysate determined. From these results, and by their comparison with those obtained previously from studies on calf-thymus deoxyribonucleic acid, it appears that either the chemical degradation proceeds differently in the two nucleic acids, or the nucleotides in one or both of the nucleic acids are not distributed randomly.

Previous papers [1] have reported a chemical method for the specific degradation of calf-thymus deoxyribonucleic acid, in which the nucleic acid was treated with mercaptoacetic acid in the presence of zinc chloride and anhydrous sodium sulphate, and the resulting *aldehydo*apurinic acid di(carboxymethyl) dithioacetal degraded at pH 12 at 37°. This cleaved the phosphodiester linkages, which in the original deoxyribonucleic acid were to purine nucleotides, but left mainly intact the linkages between pyrimidine nucleotides. The results showed that the pyrimidine nucleotides occurred singly, in pairs, and in groups of three, but that a simple uniform arrangement, such as alternate purines and pyrimidines did not occur.

The deoxyribonucleic acid of *Mycobacterium phlei* has now been subjected to the same procedure. Comparison of the results obtained for the two nucleic acids may indicate whether or not, in deoxypentose nucleic acids, the nucleotides are arranged randomly, as has been discussed in the case of pentosenucleic acids by Gamow and Ycas.[2]

The deoxyribonucleic acid was isolated from *Mycobacterium phlei* as previously described, protein being removed by the use of chloral hydrate.[3] Difficulty was experienced in completely removing the ribonucleic acid which contaminated the product. Fractionation by means of the " Cetrimide " salts, or by adsorption on charcoal,[4] which had been successful in previous cases, did not give a satisfactory product. It appeared

* Part III, *J.*, 1956, 2584.

1 Jones and Letham, *J.*, 1956, 2573; Jones, Letham, and Stacey, *ibid.*, pp. 2579, 2584.
2 Gamow and Ycas, *Proc. Nat. Acad. Sci. U.S.A.*, 1955, **41**, 1011; Gamow, *Sci. American*, 1955, **192**, No. 4, p. 70.
3 Jones and Marsh, *Biochim. Biophys. Acta*, 1954, **14**, 559.
4 Dutta, Jones, and Stacey, *ibid.*, 1953, **10**, 613.

from this and other work that, when the nucleic acids are isolated under mild conditions, the ribonucleic acid remains relatively undegraded and, therefore, resembles closely the deoxyribonucleic acid in physical properties, thus making difficult the separation of the two types of nucleic acid. Final traces of ribonucleic acid were removed by treatment with crystalline ribonuclease (free from deoxyribonuclease) followed by fractionation with "Cetrimide." The deoxyribonucleic acid contained more guanine and cytosine than adenine and thymine. The low recovery of purines and pyrimidines with respect to phosphorus was probably due to the presence of a phosphorus-containing impurity (*e.g.*, a polyphosphate).

The reaction with mercaptoacetic acid was carried out by a slight modification of the previous method.[1] During the reaction, the purines were completely liberated, about 12% of the phosphorus was rendered dialysable, and the resulting dithioacetal contained the theoretical amount of sulphur. The products of the alkaline degradation of this dithioacetal were separated into 20 components by paper chromatography in propan-2-ol–ammonia and by paper electrophoresis at pH 3·5. Owing to the comparatively small amount of deoxyribonucleic acid available, and to the small proportion of some of the components present, sufficient material could not be obtained for complete pyrimidine, phosphorus, and sulphur analyses on all of the components. Some were identified by comparison with similar substances identified in hydrolysates of calf-thymus *aldehydo*-apurinic acid di(carboxymethyl) dithioacetal. For the estimation of cytosine and thymine, a spectrophotometric method was used, which gave results within 10% of those obtained by the usual hydrolysis followed by paper chromatography and could be applied to the identification of the small homogeneous components. The amounts of each component present were estimated by cutting the paper chromatograms and electrophoretograms into strips, eluting these, and measuring the optical density at 276 mμ. This method accounts for about 80% of the ultraviolet-absorbing material of the hydrolysate.

TABLE 1. *Components of the alkaline hydrolysate of the* aldehydo*apurinic acid di(carboxymethyl) dithioacetal of* Mycobacterium phlei *and of calf thymus.*

	% of the total pyrimidines				
	Detd. in dithioacetal from		Calculated ‡ (on basis of random distribution)		
Component †	*Myco. phlei* *	Calf thymus *	Sequence of pyrimidines	*Myco. phlei*	Calf thymus
T	2·0	3·0	t	7·8	13·8
C	4·0	2·0	c	17·3	11·3
T–P–C	3·0	4·5	tc	10·6	12·4
T–P–T	0	2·0	tt	2·5	7·6
C–P–C	1·5	1·5	cc	11·9	5·1
T–P–S–P	0·5	9·0	Trinucleotides + larger oligonucleotides	50·0	50·0
C–P–S–P	2·0	4·5			
T–P–T–P–S–P	?	5·0			
T–P–C–P–S–P	1·0	10·0			
T–P–T–P–C–P–S–P	1·0	3·0			
Zone 1	62·5	45·0			

* To the nearest 0·5%. † T = thymidine residue; C = deoxycytidine residue; S = *aldehydo*-2-deoxy-D-ribose, di(carboxymethyl) dithioacetal residue; P = phosphate residue. The sequence in each component has not been determined. ‡ See text.

Results are in Table 1. The "calculated" figures are derived by the following mathematical treatment, for which we are indebted to Dr. A. R. Peacocke.

Consider a large number of linear arrays each containing $N/2$ purines and $N/2$ pyrimidines, in which each site has an equal chance of being occupied by a purine or a pyrimidine. This latter condition corresponds to the observation that in deoxyribonucleic acid

$$[\text{adenine}] + [\text{guanine}] = [\text{cytosine}] + [\text{thymine}]$$

The number of sequences of n pyrimidines will be $N/2^{n+2}$, *i.e.*, the number of sequences of 1, 2, and 3 pyrimidines will be $N/2^3$, $N/2^4$, and $N/2^5$ respectively.

The total number of pyrimidines occurring in all unbroken sequences of length n will be $Nn/2^{n+2}$. If it is assumed that $N \gg n$ and that there is a very large number of arrays of N units, the total number of pyrimidines in a given array of N will be on the average $N \sum_{0}^{\infty} n/2^{n+2}, = \frac{1}{2}N$, as at first assumed.

Hence the fraction of pyrimidines occurring singly ($n = 1$) $= N/2^3 \div N/2 = \frac{1}{4}$, and the fraction occurring in pairs ($n = 2$) $= 2N/2^4 \div N/2 = \frac{1}{4}$. Therefore the fraction of pyrimidines occurring in sequences greater than two ($n > 2$) $= \frac{1}{2}$.

Let the fraction of pyrimidines which are cytosine (c) and thymine (t) be c and t respectively. Then in deoxyribonucleic acid $c + t = 1$. For single units of pyrimidines, the fraction of pyrimidines occurring as cytosine will be $c/4$ and those occurring as thymine will be $t/4$.

For units of two pyrimidines, $c^2/4$ will occur as cc, $2ct/4$ as ct, and $t^2/4$ as tt.

For units containing n nucleotides the proportions of the various arrangements can be obtained by expansion of $(c + t)^n = 1$.

Although many of the components from both *Myco. phlei* and calf-thymus deoxyribonucleic acids have not been identified, certain differences between the two nucleic acids can be seen (Table 1). Thus, the proportion of pyrimidines in zone 1 (which probably contains mainly trinucleotides and larger oligonucleotides) is much higher from *Myco. phlei* deoxyribonucleic acid than from calf-thymus deoxyribonucleic acid. This makes the proportion of the other components correspondingly lower, except that of the mononucleosides which are present in similar amounts in the two cases. The explanation of this result cannot be that cytosine-containing components have lower R_F values than thymine-containing ones, since this is not so with mono- and di-nucleotide derivatives. If it is assumed (*a*) that the pyrimidines in deoxyribonucleic acid are distributed randomly, (*b*) that the non-specific degradation of the deoxyribonucleic acid which may occur on treatment with mercaptoacetic acid is not influenced by the type of pyrimidine nucleotide undergoing fission, and (*c*) that the mode of fission occurring during the alkaline degradation of the *aldehydo*apurinic acid di(carboxymethyl) dithioacetal is also not influenced by the type of nucleotide, then it would be expected that the ratio of the amount of one component obtained from one deoxyribonucleic acid to that of the same component from a second deoxyribonucleic acid would be the same as the ratio of the calculated probabilities of the occurrence of the particular sequence in the two nucleic acids. The results (Table 1) show that this is not the case, the most significant deviations being with the components, T-P-C-P-S-P and C-P-C. In the hydrolysate of the dithioacetal from one deoxyribonucleic acid, if the above conditions held, the proportions of the amounts of the dinucleoside phosphates to each other would be proportional to the probability of the occurrence of the particular sequence. This is approximately so with the dinucleoside phosphates from calf-thymus deoxyribonucleic acid, but not with those from *Myco. phlei* deoxyribonucleic acid, the amount of di(deoxycytidine) phosphate being only half that which would be expected from the amount of thymidine-deoxycytidine phosphate found.

These results indicate that one or more of the conditions outlined above do not apply. With respect to (*b*), the amount of dialysable phosphorus produced during formation of the dithioacetal should give a measure of the non-specific degradation. Calf-thymus deoxyribonucleic acid gave 6% and *Myco. phlei* deoxyribonucleic acid gave 12% of dialysable phosphorus. This difference fails to account for the observed differences in the amount of zone 1. Furthermore, in recent experiments with calf-thymus deoxyribonucleic acid in which 12% of dialysable phosphorus was produced upon conversion into the dithioacetal, again about 45% of zone 1 was obtained. These results suggest that in one or both of these deoxyribonucleic acids, the nucleotides are not randomly distributed, although the possibility must not be overlooked that the type of pyrimidine influences the mode of

breakdown of the nucleic acid. A definite conclusion on this point should be made possible by the complete identification of all the components from both deoxyribonucleic acids.

Experimental

Nitrogen was determined by the method of Ma and Zuazaga,[5] phosphorus by that of Jones, Lee, and Peacocke,[6] and sulphur by that of Jones and Letham.[7] The purine and pyrimidine contents of *Myco. phlei* deoxyribonucleic acid and the corresponding *aldehydo*apurinic acid di(carboxymethyl) dithioacetal were determined by hydrolysis with anhydrous trifluoroacetic acid at 155° for 50 min. This reagent gave a slightly better recovery of thymine than did formic acid. The purines and pyrimidines were separated by paper chromatography in propan-2-ol–hydrochloric acid [8] and estimated spectrophotometrically.

Phosphomonoesterase was prepared from human prostate gland by Markham and Smith's method.[9] Previously (Part III) this enzyme was state to have ribonuclease activity. It has now been shown that this was not the case, the previous results possibly being due to the presence of enzyme in the ribonucleic acid used for the estimation.

Isolation of Mycobacterium phlei *Deoxyribonucleic Acid.*—*Mycobacterium phlei* Sõhn was grown on a medium (5 l.) containing Lab. Lemco 1%, Oxoid bacteriological peptone 1%, sodium chloride 0·5%, and glycerol 5%, distributed in portions (500 ml.) in mould-culture flasks. The cells were harvested after 14 days' growth at 37° and the nucleic acids isolated by Jones and Marsh's method [3] with the use of chloral hydrate to remove most of the protein. The product thus isolated (1·52 g.) contained 20—40% of ribonucleic acid. Attempts to remove this by fractionation of the " Cetrimide " salts or by adsorption on charcoal [4] were not completely successful although both methods substantially reduced the amount of ribonucleic acid present. After treatment by these two methods, the remaining ribonucleic acid was degraded with crystalline pancreatic ribonuclease (free from deoxyribonuclease), the enzyme was removed as a gel by shaking the solution with chloroform and octanol (9 : 1), and the deoxyribonucleic acid in the aqueous phase was purified *via* the " Cetrimide " salt.[4] The material finally isolated had the following composition :

Nitrogen (%)	Phosphorus (%)	Moles of base/4 g.-atoms of phosphorus				
		Guanine	Cytosine	Adenine	Thymine	Total
14·6	8·8	1·31	1·30	0·60	0·59	3·80

No ribonucleic acid could be detected. The recovery of bases relative to phosphorus was lower than the theoretical value. This may have been due to the presence in this nucleic acid of a small amount of a polyphosphate, the presence of which was indicated by the formation of a small, granular, phosphorus-containing precipitate on the addition of barium chloride at pH 7.

Mycobacterium phlei aldehydo*Apurinic Acid Di*(*carboxymethyl*) *Dithioacetal.*—Preliminary experiments indicated that the reaction with mercaptoacetic proceeded better with deoxyribonucleic acid which had been dried *in vacuo* over phosphoric anhydride at 110° for 90 min. than with that which had been more extensively dried (100° *in vacuo* over phosphoric anhydride for 12 hr.). In the latter case, only 90% of the theoretical amount of sulphur was incorporated into the product, although there was complete liberation of guanine.

Myco. phlei deoxyribonucleic acid (340 mg. dry wt.), dissolved in 98% mercaptoacetic acid (8·8 ml.), was treated with a solution of anhydrous zinc chloride (1·98 g.) in mercaptoacetic acid (5·5 ml.), containing in suspension anhydrous sodium sulphate (1·06 g.), for 24 hr. at 37°, the whole being shaken at frequent intervals. Ether (8 vols.) was then added, and the resulting precipitate centrifuged off, washed with ether (3 times), and shaken overnight with a solution of glycine (2 g.) in acetate buffer (pH 6·0; 40 ml.). The resulting suspension was dialysed for a total of 6 days against two portions (2 × 1 l.) of 0·5% glycine solution and then centrifuged. The precipitate was extracted with 5% glycine solution and the extract dialysed against 1% glycine solution. The supernatant liquid and the 5% glycine extract were combined and

[5] Ma and Zuazaga, *Ind. Eng. Chem. Anal.*, 1942, **14**, 280.
[6] Jones, Lee, and Peacocke, *J.*, 1951, 623.
[7] Jones and Letham, *Analyst*, 1956, **81**, 15.
[8] Wyatt, *Biochem. J.*, 1951, **48**, 584.
[9] Markham and Smith, *Biochem. J.*, 1952, **52**, 558, 565.

centrifuged and the supernatant liquid dialysed exhaustively against tap-water and finally against distilled water. Freeze-drying the solution gave the dithioacetal as a white flocculent solid (280 mg.) which contained no purines and 1·30 moles of cytosine, 0·59 mole of thymine, and 3·80 g.-atoms of sulphur per 4 g.-atoms of phosphorus. No free aldehyde groups could be detected.

The ethereal mother-liquors contained about 2% of the cytosine which was originally present in the deoxyribonucleic acid, but no thymine. During the isolation 12% of the total phosphorus diffused through the Cellophane membranes.

Alkaline Degradation of the Dithioacetal.—The dithioacetal (40 mg.) was dissolved in sodium hydroxide (pH 12; 30 ml.) and kept at this pH for 26 hr. at 37° (dilute sodium hydroxide being added at intervals as necessary). The solution was then neutralised with dilute acetic acid and dialysed. The amount of ultraviolet-absorbing material diffusing through the Cellophane membrane was measured by diluting samples with 0·1N-hydrochloric acid and measuring the optical density at 268 mμ (the wavelength at which the molecular extinctions of thymidylic acid and deoxycytidylic acid coincide at pH 1). When the dialysis was conducted against distilled water, 81% of the material diffused through the membrane and the remainder diffused through upon dialysis against M-sodium chloride.

The neutralised alkaline hydrolysate was separated by chromatography on Whatman No. 3 paper in propan-2-ol–water–ammonia (*d* 0·880) (70 : 30 : 6) into 10 zones. The % of the pyrimidines in each zone was determined as described previously [1] and the total equated to 100%. The properties of the various zones were as follows:

Zone	1	2	3	4	5	6	7	8	9	10
R_F	—	—	—	—	—	0·15	0·24	0·42	0·60	0·67
R_C *	0·07	0·15	0·24	0·37	0·69	1·20	1·60	—	—	—
% of total pyrimidine	62·5	8·3	5·2	3·1	5·3	1·8	3·8	4·4	3·8	2·0

* R_C = distance moved by zone/distance moved by cytidylic acid.

The components of the various zones were prepared by elution of a number of chromatograms. They were then run in the same solvent system before and after re-hydrolysis at pH 12. In neither case was there a significant change in R_C value. The components of zones 2, 3, 4, and 8 were then separated by paper electrophoresis at pH 3·5 into a total of 14 fractions the properties of which are shown in Table 2.

Pyrimidine Analyses.—Owing to the small quantities of these fractions available, it was not possible to determine the pyrimidine contents in the usual way. A method was developed, therefore, in which the pyrimidines were determined spectrophotometrically without prior hydrolysis or chromatographic separation. This depended on the fact that, at pH 1, the maximum optical absorption of thymidylic acid is at 267 mμ and that of deoxycytidylic acid is at 280 mμ. Hence, from measurements at these two wavelengths, the proportions of each of these nucleotides present in a mixture of the two could be calculated. When applied to zones of known composition, the method gave results which were within 10% of that expected. The error was probably due to interaction between the bases in the di- and the tri-nucleotide components. Despite this error, however, the method could be used as a semiquantitative measure of the base content of the smaller components and, when taken with other evidence, could be used in their identification. The results are recorded in Table 2.

Identification of the Components.—*Zone* 10. This zone was not further resolved by any method tried. It contained thymine and 2-deoxy-D-ribose residues, but no cytosine, sulphur, or phosphorus. It was identified as thymidine by comparison with authentic material.

Zone 9. This was similar to Zone 10 but contained cytosine instead of thymine. It was identified as deoxycytidine by comparison with authentic material.

Zone 8. This sulphur-free zone was separated by paper electrophoresis at pH 3·5 into two components, 8A and 8B. The pyrimidine contents of these were estimated by the spectrophotometric method. 8A contained 1·0 mole of thymine and 0·8 mole of cytosine per g.-atom of phosphorus, and 8B contained 2·2 moles of cytosine per g.-atom of phosphorus. Treatment of 8A and 8B with phosphomonoesterase did not alter their R_F values in the propan-2-ol–ammonia solvent. This showed the absence of a terminal phosphate group. The identification as dinucleoside phosphates as indicated in Table 2 was consistent with their electrophoretic behaviour. No dithymidine phosphate was detected, although from previous work with calf-thymus deoxyribonucleic acid,[1] it would be expected to occur in zone 8.

TABLE 2. *Properties of the components of the alkaline hydrolysate of* Myco. phlei aldehydo*apurinic acid di(carboxymethyl) dithioacetal.*

Zone	% of total pyrimidines	Thymine / Cytosine (moles) †	Proposed structure *
1	62·5	0·50	—
2A	1·0	1·9	T–P–T–P–C–S–P
2B	1·4	1·1	C–P–T–P–S–P
2C	1·8	0·60	—
2D	1·9	0·50	—
2E	1·2	0·60	—
2F	1·0	—	—
3A	2·7	1·0	—
3B	1·6	Cytosine only	—
3C	0·8	Thymine only	—
4A	0·6	Thymine only	T–P–S–P
4B	2·1	Cytosine only	C–P–S–P
4C	0·4	1·2	—
5	5·3	—	—
6	1·8	—	—
7	3·8	—	—
8A	2·9	1·2	C–P–T
8B	1·5	Cytosine only	C–P–C
9	3·8	Cytosine only	C
10	2·0	Thymine only	T

* T = Thymidine residue; C = deoxycytidine residue; S = *aldehydo*-2-deoxy-D-ribose di(carboxymethyl) dithioacetal residue; P = phosphate residue. The sequences within zones 2A and 2B have not been determined, these being only possibilities. † Semiquantitative determination.

Zone 4. This zone separated into components 4A, 4B, and 4C, on electrophoresis at pH 3·5. The identification of 4A and 4B, as indicated in Table 2, was made from the base contents and by comparison of their behaviour upon paper chromatography and paper electrophoresis with components already identified in previous work on calf-thymus deoxyribonucleic acid. These results are summarised as follows:

Component	R_C	M_C †	Proposed structure *
Myco. phlei 4A	0·37	9·3	T–P–S–P
Calf-thymus 3A	0·35	9·1	T–P–S–P
Myco. phlei 4B	0·37	6·8	C–P–S–P
Calf-thymus 3C	0·35	6·6	C–P–S–P

* See Table 2. † M_C = Distance moved by component relative to distance moved by cytidylic acid (corrected for endosmosis).

Zone 2. This zone separated into 6 components by paper electrophoresis at pH 3·5. Components 2A and 2B corresponded in base content and R_C value (0·15) in propan-2-ol–ammonia with components from calf-thymus deoxyribonucleic acid identified as indicated in Table 2. Their behaviour on electrophoresis was consistent with these structures.

Zone 1. This was heterogeneous. No definite resolution was obtained either by two-dimensional paper electrophoresis or by electrophoresis in the Antweiler micro-electrophoresis apparatus.

The authors thank Mrs. S. S. Topaz and Mr. E. T. J. Chelton, F.I.M.L.T., for technical assistance. One of us (M. S.) thanks the Medical Research Council for a grant for personal assistance by B. E. W.

CHEMISTRY DEPARTMENT, UNIVERSITY OF BIRMINGHAM,
EDGBASTON, BIRMINGHAM, 15.
[*Received, January 22nd*, 1957.]