EVIDENCE OF COMMON REGULARITIES IN THE COMPOSITION OF PENTOSE NUCLEIC ACIDS*

by

DAVID ELSON AND ERWIN CHARGAFF

Cell Chemistry Laboratory, Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York (U.S.A.)

General regularities governing the composition of deoxypentose nucleic acids (DNA) have been known for several years[1]. Though the composition of DNA specimens may vary considerably with their source, purified preparations of widely divergent origin have been found to contain equimolar quantities of adenine and thymine and of guanine and cytosine. Consequently, the molar sum of purines equals that of pyrimidines, and for every base carrying an amino group in the 6 position (adenine, cytosine) there occurs one with a 6-keto (oxo) group (guanine, thymine).

No comparable composition regularities have emerged from the analyses of purified specimens of pentose nucleic acids (PNA) of which there is no lack in the literature[2]. In recent investigations of PNA composition in varous biological preparations[3–6] we have resorted to a procedure in which extracts of the total PNA were subjected directly to nucleotide analysis without further purification. The results indicated that, with uracil taking the place of thymine, the regularities characteristic of DNA were approximated in some, but not all, PNA extracts. One regularity was, however, found in all instances, namely, that the number of nitrogenous constituents carrying a 6-amino group was close, and on the average equal, to that of 6-keto compounds[7].

The results previously presented in a preliminary manner[7] have now been confirmed in additional studies, mainly of preparations of microbial origin. In this paper are assembled all the data thus far obtained in this laboratory that bear on the composition of the total PNA of various biological materials, together with a brief discussion of the conclusions suggested by these findings.

EXPERIMENTAL

Analytical

Full details of the procedures have been published previously[5]. After extraction with cold 7% trichloroacetic acid, ethanol, and ethanol-ether, the material was treated with NaOH at pH 13 to 13.5 and 30° for 16 to 18 hours and centrifuged. DNA was removed by precipitation at pH 3 to 3.5; the pentose mononucleotides, produced by the alkaline hydrolysis of PNA, remained in the supernatant fluid and were estimated by means of paper chromatography and ultraviolet spectrophotometry. As pointed out before[5], an excessive spread in the sextuple readings, at any of the specified wave lengths, *i.e.* a coefficient of variation above 5, disqualified an analysis.

* This work was aided by research grants from the National Cancer Institute, National Institutes of Health, United States Public Health Service, and from the Rockefeller Foundation.

References p. 375/376.

Material

The cellular and subcellular preparations that were analyzed are listed fully in Table I. Brief descriptions follow.

Vertebrate tissues and cell fractions. The preparations of isolated nuclei and cytoplasmic fractions have been described in the preceding paper[6]. Preparations designated as "whole tissue" of the ox represented portions of the lyophilized material that also served for the isolation of nuclei in non-aqueous media[6]. Whole rat kidney was analyzed as a pulp suspended in 0.25 *M* sucrose. The sediment (600 × *g*, 10 minutes) from the suspension of a similar pulp of rat liver yielded the nuclear fraction; it consisted of nuclei contaminated with unbroken cells.

Sea urchin eggs and embryos. These included specimens of embryonic stages of *Paracentrotus lividus* ranging from the unfertilized egg to the 48-hour embryo, both normal and cultured in the presence of Li^+ ions. Full details have been published[5].

Microorganisms. Yeast was either used directly as washed commercial bakers' yeast or as the harvest from freshly prepared cultures. *Saccharomyces cerevisiae* was grown at room temperature for 19 hours in a medium containing glucose, yeast extract, sodium lactate, and salts.

Serratia marcescens was cultured for 40 hours at 37° with magnetic stirring either in nutrient broth or in a medium containing 2% yeast extract, 1 to 1.5% glucose, and 0.1% asparagine.

Escherichia coli, designated as strain II in a recent publication[8], was employed in the form of 24-hour cultures at 37° in nutrient broth variously fortified with glucose (0.3 to 4%) and sometimes with peptone (1 to 2%). Both standing and shaken cultures were used.

Mycobacterium phlei (American Type Culture Collection No. 10142) was cultured at 37° in nutrient broth containing 2% glycerol and 0.05% Tween 20[9], in one case for 65 hours on a horizontally rotating platform, in another, for 3 weeks as a standing culture followed by agitation for 20 hours.

PNA analyses were carried out on (a) whole cells; (b) cell fractions, prepared by the differential centrifugation of extracts of crushed cells; (c) fractions afforded by the stepwise precipitation with ammonium sulfate from cell extracts; (d) several other extracts and residues obtained during the processing of *M. phlei*. All operations were, unless stated otherwise, performed at 0 to 3°. Yeast was ground as a paste in 0.25 *M* sucrose for 90 minutes in a bacterial mill[10]. For the rupture of the other bacteria fine glass beads ("ballotini", 0.3 mm diameter) were used; they were added in sufficient quantity, so as entirely to fill the suspension of the washed cells in cold 0.25 *M* sucrose, and the mixture was subjected to vigorous vibration for 90 minutes. Cell breakage was extensive, but not complete. The ground material was washed off the beads with 0.25 *M* sucrose and the suspension first centrifuged at 20,000 × *g* for 10 to 20 minutes; the supernatant fluid was then subjected to centrifugation at 78,000 × *g* for 60 to 75 minutes, whereby the sedimentable fraction P-78 (Nos. 14, 18, 25 in Table I) and the supernatant fraction S-78 (Nos. 12, 15, 19, 26 in Table I) were obtained.

Several of the bacterial S-78 fractions were examined further in the form of subfractions obtained by successive precipitation with increasing concentrations of ammonium sulfate over a saturation range of 20 to 100%. A total of 23 such nucleoprotein fractions was analyzed. Details of these studies are reserved for a later occasion, since for the purpose of the present discussion it was deemed sufficient to list averages rather than the analysis of each individual subfraction. Instead, the results were averaged for each bacterial species and so entered as "S-78, $(NH_4)_2SO_4$ fractions" (Nos. 16, 20, 27 in Table I).

Three additional PNA preparations from *M. phlei* were also examined. In view of the reported accumulation of extracellular nucleic acid in cultures of mycobacteria[11], the cell-free medium of a 65-hour culture was treated with one volume of ethanol, brought to pH 3, and the resulting precipitate, which contained PNA, collected by centrifugation (No. 22 in Table I). In addition, a 3 week old culture was treated by a procedure somewhat similar to that of KHOUVINE *et al.*[12]. The cells, first extracted at room temperature for 24 hours with 5 portions of ethanol-ether (1:1 by volume), for 60 hours with chloroform, and dried *in vacuo* over paraffin and H_2SO_4, were suspended at 0 to 3° in 0.14 *M* NaCl and shaken with "ballotini" for 9 hours. The PNA fraction thus removed is listed as No. 23 in Table I and the extraction residue, collected after an additional treatment with *M* NaCl, as No. 24. The PNA fraction extracted by *M* NaCl gave an unsatisfactory analysis.

RESULTS AND INTERPRETATIONS

In the presentation and discussion of the results, the following notation will be used for the sake of brevity: A will denote adenine, G guanine, C cytosine, U uracil, T thymine, or the corresponding nucleotides; Pu will stand for the purines (A + G), Py for the pyrimidines (C + U in PNA, C + T in DNA), 6-Am for the bases having an amino group in the 6 position (A + C), 6-K for those with a 6-keto group (G + U and G + T, respectively).

References p. 375/376.

TABLE I

PENTOSE NUCLEIC ACIDS: NUCLEOTIDE COMPOSITION AND MOLAR RELATIONSHIPS*

No.	Preparations analyzed	Number of preparations	Nucleotide composition (as moles per 100 moles nucleotide)				Molar ratios					
			A	G	C	U	A/U	G/C	Pu/Py	$\frac{A+U}{G-C}$	6-Am / 6-K	A/U / G/C
	Vertebrates											
	Whole tissue											
1	Ox liver	1	17.1	27.3	33.9	21.7	0.79	0.80	0.80	0.63	1.04	0.99
2	Ox kidney	1	19.7	26.7	33.4	20.2	0.97	0.80	0.87	0.66	1.13	1.21
3	Rat kidney	1	19.4	29.5	30.7	20.4	0.95	0.96	0.96	0.66	1.00	0.99
	Isolated nuclei											
4	Ox liver	2	19.65	25.75	30.10	24.50	0.805	0.865	0.830	0.795	0.990	0.940
5	Ox kidney	1	22.0	26.2	29.7	22.2	0.99	0.88	0.93	0.79	1.07	1.13
6	Rat liver	2	20.20	26.00	29.85	23.95	0.845	0.875	0.855	0.795	1.000	0.965
	Nuclear fraction											
7	Rat liver	1	18.0	21.7	33.6	26.7	0.68	0.65	0.66	0.81	1.06	1.05
	Cytoplasmic fractions											
8	Rat liver and kidney	14	17.99	31.64	30.04	20.30	0.901	1.054	0.989	0.624	0.928	0.859
9	Frog liver and kidney	5	20.18	27.24	32.08	20.50	0.990	0.856	0.906	0.688	1.098	1.174
	Sea urchin											
10	*Paracentrotus* eggs and embryos	31	22.56	29.42	27.19	20.82	1.086	1.085	1.085	0.767	0.992	1.008
	Microorganisms											
	Yeast											
11	Whole cells	3	25.40	24.60	22.63	27.37	0.930	1.087	0.997	1.120	0.927	0.867
12	S-78	1	24.1	26.0	25.1	24.8	0.98	1.04	1.01	0.95	0.97	0.94
	S. marcescens											
13	Whole cells	1	20.3	31.2	24.3	24.1	0.84	1.28	1.06	0.80	0.81	0.66
14	P-78	3	24.80	29.07	24.23	21.87	1.137	1.203	1.170	0.873	0.967	0.963
15	S-78	2	25.05	29.55	24.40	21.00	1.195	1.215	1.205	0.850	0.980	0.985
16	S-78, $(NH_4)_2SO_4$ fractions	13	25.31	28.52	25.21	20.98	1.214	1.183	1.175	0.872	1.033	1.116
	E. coli											
17	Whole cells	1	25.3	28.8	24.7	21.2	1.19	1.16	1.18	0.87	1.00	1.03
18	P-78	2	26.25	28.95	24.10	20.70	1.270	1.200	1.235	0.880	1.015	1.060
19	S-78	2	24.30	27.95	26.10	21.65	1.120	1.075	1.095	0.850	1.015	1.040
20	S-78, $(NH_4)_2SO_4$ fractions	7	24.59	29.60	25.50	20.33	1.214	1.164	1.184	0.817	1.006	1.049
	M. phlei											
21	Whole cells	2	20.85	30.75	27.10	21.25	0.980	1.130	1.065	0.730	0.925	0.865
22	Medium	1	22.6	29.9	24.9	22.6	1.00	1.20	1.10	0.83	0.90	0.83
23	0.14 *M* NaCl extract	1	21.2	32.3	27.8	18.7	1.13	1.16	1.15	0.67	0.96	0.97
24	Residue from NaCl extractions	1	21.1	32.9	26.2	19.8	1.06	1.26	1.17	0.89	0.89	0.84
25	P-78	1	23.6	30.9	24.8	20.7	1.14	1.24	1.20	0.80	0.94	0.92
26	S-78	1	22.2	30.6	26.8	20.4	1.08	1.13	1.11	0.74	0.96	0.96
27	S-78, $(NH_4)_2SO_4$ fractions	3	22.47	31.63	25.97	19.90	1.130	1.213	1.177	0.733	0.940	0.933

* Abbreviations used: A, adenylic acid; G, guanylic acid; C, cytidylic acid; U, uridylic acid; Pu, purine nucleotides; Py, pyrimidine nucleotides; 6-Am, compounds carrying an amino group in 6 position (adenylic, cytidylic acids); 6-K, compounds carrying a keto group in 6 position (guanylic, uridylic acids); P-78, fraction sedimenting at 78,000 × *g*; S-78, fraction remaining in supernatant fluid at 78,000 × *g*.

TABLE II

PENTOSE NUCLEIC ACIDS: MEAN MOLAR RELATIONSHIPS*

	Number of preparations	G/A	C/A	G/U	C/U	A/U	G/C	Pu/Py	$\frac{A+U}{G+C}$	$\frac{6\text{-}Am}{6\text{-}K}$	$\frac{A/U}{G/C}$
Vertebrates	28	1.560	1.644	1.401	1.472	0.902	0.950	0.927	0.675	0.992	0.967
Sea urchin	31	1.306	1.207	1.417	1.307	1.086	1.085	1.085	0.767	0.992	1.008
Microorganisms	45	1.217	1.049	1.379	1.192	1.143	1.171	1.151	0.853	0.984	1.004
All preparations	104	1.336	1.256	1.396	1.302	1.061	1.086	1.071	0.780	0.988	0.995
Standard error of mean	104	0.024	0.028	0.021	0.019	0.015	0.019	0.013	0.011	0.009	0.018
Coefficient of variation	104	18.4	22.4	15.6	15.2	14.3	17.5	12.2	15.0	9.1	18.2

* The abbreviations are explained in Table I. The mean values were computed from the ratios found in each individual analysis, as were the statistical terms.

All analytical data collected in this laboratory up to the present time on pentose nucleic acids that had not been subjected to special isolation and purification procedures are assembled in Table I. These include every instance in which the composition of the total PNA of a sample was determined, as outlined above, and cover 104 specimens of widely varying origin. Of the molar ratios included in this table those for A/U and G/C correspond to those having a value of 1 in DNA[1]. The ratios Pu/Py, (A + U)/(G + C), and 6-Am/6-K represent the three possible ways of pairing the bases. The expression (A/U)/(G/C) will be considered later. It will be noted that many of the values in Table I are composite values, being averages derived from several preparations.

Table II contains mean values for a series of ten molar ratios, together with certain statistical data computed from the 104 individual analyses. The first nine ratios—the significance of the last column will be considered later—exhaust the possibilities of comparing the relative amounts of single nucleotides and pairs of nucleotides. In addition, the findings have been grouped according to the phylogenetic origin of the specimens. Since two of the three groups comprise, within themselves, samples of quite different composition (compare Table I), no particular biological significance can be attached to this arrangement; it merely demonstrates that pentose nucleic acids of different composition exist in nature. Of these nine ratios, eight differ from one group to the next. The 6-Am/6-K values, however, are identical and virtually equal to unity. Of the single means calculated for all 104 samples, the mean 6-Am/6-K ratio, with a value very close to 1, shows the lowest coefficient of variation (standard deviation as percentage of the mean); *i.e.* the individual values are arranged most closely around the mean.* Moreover, the frequency distribution of the individual values for this ratio, shown in Fig. 1, approximates that of a normal curve of error. The most characteristic and least varying feature of PNA composition demonstrated by these results appears, therefore, to be the presence of nucleotides with 6-amino and with 6-keto groups in virtually equal number.

A second regularity is suggested by Table I, namely, that the ratios A/U and G/C (and therefore Pu/Py) tend to equal each other, or nearly so. In certain instances these ratios are, in addition, very close to 1 (compare Table 1 of a previous publication[7]), as

* If the three extreme values found for the 6-Am/6-K ratio (0.7, 1.3, 1.3) are omitted, the coefficient of variation for the remaining 101 specimens is reduced from 9.1 to 7.3.

References p. 375/376.

is generally true of DNA, with thymine taking the place of uracil[1]. Continued analytical studies on PNA, mainly of microbial preparations, have, in general, confirmed the equality of these ratios, but have indicated that their value is not necessarily 1. The validity of the relationship A/U = G/C can best be tested by an examination of the ratio (A/U)/(G/C). For this reason, this expression is listed in the last columns of Tables I and II. It will be seen (Table II) that the mean value for this ratio is very close to 1.

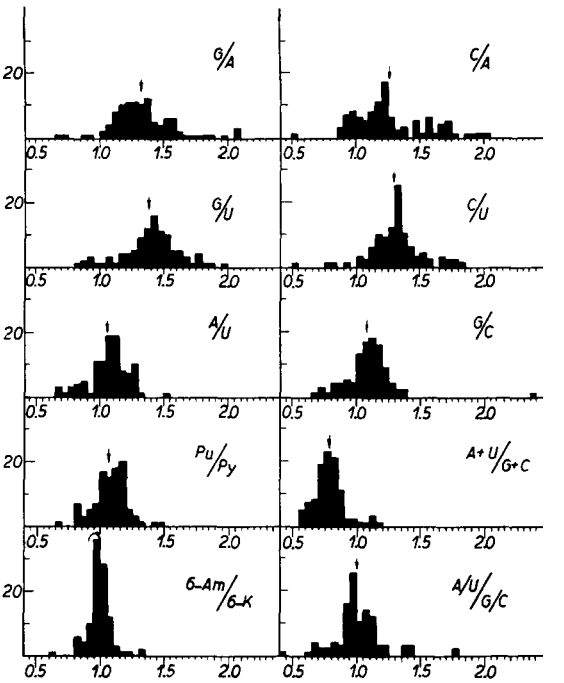

Fig. 1. Frequency distributions of nucleotide ratios in 104 PNA analyses. Ordinates: number of times value occurred. Abscissae: value of the ratio (each division represents an increment of 0.05). Arrows indicate means. For meaning of symbols, see text.

It can, however, be easily shown mathematically that if the sum of A and C equals that of G and U, *i.e.* if 6-Am = 6-K, the relation A/U = G/C can hold true only in two special cases; namely, (a) if A = G and U = C; or (b) if A = U and G = C. The proposition (a) is excluded by the evidence of Table I. Though this is less obvious, it appears that proposition (b) also does not hold, in a general fashion, for our results. If (b) were true, the ratios A/U and G/C would equal 1. As pointed out above, they do so only in a limited number of cases. It appears, in fact, that the probability of all values of A/U and G/C being unity is less than 0.01, since for both ratios the value of 1.00 lies outside the range of the expression (mean $\pm$ 3 $\times$ standard error) (see Table II). Thus, it is unlikely that the relationships A = U and G = C are generally characteristic of PNA composition.

It is, therefore, necessary to decide between the two composition regularities indicated by Table II, for mathematically they are mutually exclusive. Of these relationships, namely, A/U = G/C (or (A/U)/(G/C) = 1) and A + C = G + U (or 6-Am/6-K = 1), the latter is by far the more likely choice. The decision in favor of 6-Am = 6-K is made on two grounds: the comparison of the coefficients of variation (Table II) and of the frequency distributions (Fig. 1). The coefficients of variation show the scatter of the individual values to be much smaller for 6-Am/6-K than for (A/U)/(G/C). The frequency distributions for the several expressions listed in Table II are plotted in Fig. 1. It can be seen that the histogram for 6-Am/6-K approximates a normal curve of error, showing random scattering about a well defined maximum which coincides with the mean. On the other hand, the histogram for (A/U)/(G/C), showing two maxima, indicates that the individual deviations from the mean are not due solely to fortuitous experimental error and suggests that the true value of this ratio was not the same in all of the samples studied.

The literature contains a not inconsiderable amount of information about the

References p. 375/376.

nucleotide composition of PNA preparations that had not undergone purification procedures that could have resulted in fractionation. In the following cases the PNA analyzed was probably the total PNA of the sample or representative of it: mammalian whole tissues and cell fractions[13-20], fowl nuclei[15], echinoderm whole eggs [21] and oöcyte fractions[22], yeast[23], plant viruses[24-29], and pentose nucleoproteins or extracts containing pentose nucleoproteins[12, 30-34]. These studies present data on PNA composition derived from 151 specimens. Of these, 91 show a 6-Am/6-K ratio of between 0.90 and 1.10. The mean value for all 151 samples is 0.93. This is considerably lower than the value which we have reported here; and a survey of the literature, unsupported by the evidence presented in this paper, probably would not have led to the conclusion that 6-Am equaled 6-K in PNA. It may, however, be of some significance that in the majority of total PNA analyses reported by other workers the value of the 6-Am/6-K ratio was quite close to unity.

It should be pointed out here that among the plant viruses there appear to be at least two cases in which 6-Am does not equal 6-K, namely, turnip yellow mosaic virus[25] and potato virus X[28].

A similar survey of the literature on the nucleotide composition of purified PNA preparations[2] would be of little value in the present context. Considerable fluctuations in composition often are observed, even in specimens of the same origin. As an illustration it may be mentioned that in seven PNA preparations from yeast isolated and analyzed in this laboratory values for 6-Am/6-K ranging from 0.62 to 0.94, with a mean value of 0.75, were recorded.

DISCUSSION

The principal purpose of this inquiry has been the attempt to ascertain whether there exist features that are shared by all pentose nucleic acids. We believe that one such common regularity has been established for the preparations considered in this paper, namely, the equality of the number of nucleotides carrying an amino group in the 6 position and of those having a 6-keto (or 6-oxo) group; and that, in view of the diverse origin of these preparations, the suggestion is warranted that this relationship may be general. It is clear that great caution must be observed in proposing general regularities for such a complex, unstable and poorly understood macromolecular species as PNA. Even so, the results compiled in Table I, though less plentiful, and perhaps less convincing, than the evidence that has led to the recognition of composition regularities in DNA[1], show the relationship 6-Am = 6-K to be maintained with remarkable consistency in the total PNA of cellular material of widely divergent origin; this regularity has persisted in subfractions prepared by several different, mild procedures and has emerged more clearly as the number of analyses has increased.

Assuming that this generalization regarding PNA composition is correct, we may now proceed to consider its implications. A limited resemblance to DNA is obvious. In the latter group of substances the experimentally observed regularities may be summarized as follows: A/T = G/C = Pu/Py = 6-Am/6-K = 1. It is the last-mentioned relationship, 6-Am = 6-K, that PNA and DNA appear to have in common.

As regards DNA, a model has been proposed recently[35] on the basis of X-ray diffraction data[36-38] and the unity relationships or the specific pairing of bases which had been known for some time[39]. This model proposes a helix composed of two parallel strands

References p. 375/376.

which turn about a common axis and are held together by hydrogen bonds between specific pairs of bases, one in each strand. The base pairs are linked by two hydrogen bonds, one between the 1 positions, the other between the 6-substituents of the rings. The hydrogen bonds envisioned in this scheme can form only if one base carries a 6-amino group and the other a 6-keto group (6-Am = 6-K). To maintain parallelism, a purine must bond with a pyrimidine (Pu = Py). From this it follows that A must equal T and G must equal C.

This suggestive, though unproved and in some features probably incomplete[40,41], hypothesis contains a plausible explanation of the relationship 6-Am = 6-K, *viz.*, that it is imposed by specific and regular hydrogen-bonding between the groups in question. In applying similar considerations to PNA, two alternatives are apparent: (a) a variant of the structural model proposed for DNA in which the 6-amino and the 6-keto constituents of PNA are bonded to each other; (b) an arrangement in which these groups are linked to another structure so constituted as to bind an equal number of each. Our data do not permit us to prefer one proposition to the other.

The first of these possibilities suggests a paired structure, which is, however, not identical with the DNA model mentioned before, as in PNA the ratio Pu/Py does not consistently equal 1 or another constant. Thus, the paired strands could not be equidistant at all points. Complete parallelism could be established, however, if the PNA were a mixture of at least two double structures. The simplest instance of a PNA with Pu > Py would imply one structure, in which all the U and C are paired with equivalent amounts, respectively, of A and G, and a second structure of two parallel, all-purine chains, in which A is paired with G. With Pu < Py, the second structure would contain only pyrimidines, C being linked to U. Such an arrangement could account for all the nucleotides in any PNA in which 6-Am equaled 6-K. There exists, however, to our knowledge no evidence of the occurrence of such all-purine or all-pyrimidine components of PNA. Such double structures would differ in their dimensions from compounds composed of both purines and pyrimidines and might be expected to yield characteristic X-ray diffraction patterns. An all-purine component, if it exists, could probably be isolated owing to the resistance of such a structure to pancreatic ribonuclease[42]. A suitable starting material would be a PNA specimen with a high Pu/Py ratio, such as that described by KHOUVINE *et al.* as occurring in *M. phlei*[12]. We have, however, been unable to duplicate this finding (see Nos. 21–27 in Table I).

If, in the second alternative mentioned before, PNA is bonded not to itself, but to a different substance, this substance must contain a regularly repeating unit capable of binding one amino group for each keto group. This condition is met by the proteins and polypeptides, the carbonyl and amino groups of the peptide link being able to participate in hydrogen bonding[43].

While no detailed X-ray examination of PNA has yet been reported, it seems unlikely that the distance between adjacent bases in the polynucleotide chain would differ greatly from that in DNA. It was first pointed out by ASTBURY[44] that this dimension is equal to the distance between adjacent peptide bonds in the extended polypeptide chain. If one posits, as an extreme case, an extended polypeptide chain in which every peptide carbonyl is linked by a hydrogen bond to the 6-amino group of adenine or cytosine and every peptide amino group to the 6-keto group of guanine or uracil, there results an arrangement in which the polypeptide chain is linked to two polynucleotide chains, since each peptide bond would bind two purine or pyrimidine bases and there would be

References p. 375/376.

room along a single polynucleotide chain for only one base per amino acid residue. The peptide bond would probably have to be in the *trans* configuration, with each of its two hydrogen bonds going to a different polynucleotide chain.

Such an arrangement may suggest a template, though it would not indicate which partner acted as the template for the other. PNA may conceivably function in holding the protein in a primordial form by blocking the formation of hydrogen bonds between the amide groups of the peptide chain, so that the protein could only after the removal of the PNA assume its definitive three-dimensional configuration. These are, however, matters on which direct information would have to be furnished by a different type of investigation.

There is no record in the literature of the isolation of a pentose nucleoprotein that would fulfil the requirements of the structure proposed here. But it should be pointed out that such a nucleoprotein could easily have escaped recognition, being classified as a nucleic acid. Two hypothetical instances, based on the mean composition of a PNA derived from the data in Table II (P 9.0%), may be cited: A triad composed of two polynucleotide chains and one polyglycine chain would contain no more than 7.6% of the polypeptide and have a phosphorus content of 8.3%; even with polytryptophan, the polypeptide content would be 21.2% and the P content 7.1%. Furthermore, a complex of this type would have all its phosphoric acid groups available for reaction with additional proteins. The only ribonucleoprotein that to our knowledge has been fully characterized with respect to amino acid, purine and pyrimidine contents is that isolated by PARSONS from *Clostridium perfringens*[33]. This substance showed a 6-Am/6-K ratio of 1.02, but contained 2.2 amino acid residues per base instead of the 0.5 required by the formulation discussed above; only one quarter of the available peptide bonds could, therefore, have been fully engaged in hydrogen-bonding with the nucleic acid bases.

Further work will be required to show whether the scheme discussed in the preceding paragraphs or a variant of this scheme merits adoption. But it should, in any event, be remembered that PNA occurs in nature in close conjugation with proteins. It has been observed frequently in the course of preparation of PNA that pentose nucleoproteins resist complete deproteinization much more tenaciously than do the DNA-protein complexes. The plant viruses cannot be discussed in the present context; they are probably much more complex than the structures considered by us here. But there exists evidence that there occur in certain plant viruses PNA-protein links that are not electrostatic, and some findings point to hydrogen bonds being involved[45]. In this laboratory the relationship 6-Am = 6-K has been found only in preparations from which the protein had not been removed; it vanished as the PNA was "purified". It is conceivable that, in probable contrast to DNA, the structure of at least some pentose nucleic acids is inextricably associated with the structure of the protein with which it is conjugated[46] and that pentose nucleoproteins, cautiously isolated and suitably purified, rather than PNA may be the proper object of certain physical and chemical investigations.

ACKNOWLEDGEMENT

The authors wish to express their gratitude to Mrs. LEAH WENIG TRENT for technical assistance with many of the preparations and analyses, and to Professors HOWARD LEVENE and ALLAN BIRNBAUM of the Department of Mathematical Statistics, Columbia University, for very helpful discussions of the statistical methods used in this paper and their interpretation.

References p. 375/376.

SUMMARY

1. The results of a survey of the nucleotide composition of pentose nucleic acids (PNA) have been presented. The material analyzed included whole tissues and various subcellular fractions derived from beef, rat and frog liver and kidney, from sea urchin eggs and embryos, and from several microorganisms.

2. It appears characteristic of all preparations studied that the bases with 6-amino groups (adenine, cytosine) and those with 6-keto groups (guanine, uracil) occur in approximately equal number.

3. This regularity was found only when the total PNA of the sample was analyzed.

4. Some of the structural implications of the findings are discussed.

RÉSUMÉ

1. Les résultats d'une étude de la composition en nucléotides des acides pentosenucléiques (PNA) sont exposés. Les produits analysés comprennent des tissus entiers et diverses fractions subcellulaires isolées du foie et du rein du boeuf, du rat et de la grenouille, des oeufs et des embryons d'oursins, et de plusieurs microorganismes.

2. Une caractéristique de toutes les préparations étudiées est que les bases possédant un groupe aminé en 6 (adénine, cytosine) et celles possédant un groupe cétonique en 6 (guanine, uracile) sont présentes en nombres approximativement égaux.

3. Cette régularité ne s'observe que si l'on analyse le PNA total des échantillons.

4. Quelques conséquences structurales de ces résultats sont discutées.

ZUSAMMENFASSUNG

1. Die Arbeit gibt einen Überblick über die Nukleotidzusammensetzung einer Reihe von Pentosenukleinsäuren (PNS). Das Analysenmaterial umfasste Gewebe und zahlreiche Zellfraktionen, die aus Rinder-, Ratten- und Froschleber und -niere, aus Seeigeleiern und -embryonen und aus Mikroorganismen isoliert worden waren.

2. Eine für alle Präparate charakteristische Regelmässigkeit bestand darin, dass die 6-Aminogruppen enthaltenden Basen (Adenin, Cytosin) und die 6-Ketoverbindungen (Guanin, Uracil) in fast gleicher Zahl auftraten.

3. Diese Regelmässigkeit konnte nur dann gefunden werden, wenn die gesamte PNS analysiert wurde.

4. Einige Strukturfolgerungen, die sich aus diesen Befunden ergeben, werden diskutiert.

REFERENCES

1 E. CHARGAFF, in E. CHARGAFF AND J. N. DAVIDSON, *The Nucleic Acids*, New York, 1 (1955) 307.
2 B. MAGASANIK, in E. CHARGAFF AND J. N. DAVIDSON, *The Nucleic Acids*, New York, 1 (1955) 373.
3 D. ELSON AND E. CHARGAFF, *Federation Proc.*, 10 (1951) 180.
4 D. ELSON AND E. CHARGAFF, in W. D. MCELROY AND B. GLASS, *Phosphorus Metabolism*, Baltimore, 2 (1952) 329.
5 D. ELSON, T. GUSTAFSON AND E. CHARGAFF, *J. Biol. Chem.*, 209 (1954) 285.
6 D. ELSON, L. W. TRENT AND E. CHARGAFF, *Biochim. Biophys. Acta*, 17 (1955) 362.
7 D. ELSON AND E. CHARGAFF, *Nature*, 173 (1954) 1037.
8 S. ZAMENHOF AND G. GRIBOFF, *Nature*, 174 (1954) 306.
9 R. J. DUBOS, *Proc. Soc. Exp. Biol. Med.*, 58 (1945) 361.
10 V. H. BOOTH AND D. E. GREEN, *Biochem. J.*, 32 (1938) 855.
11 F. B. SEIBERT, *Chem. Rev.*, 34 (1944) 107.
12 Y. KHOUVINE, M. BARBIER AND L. WYSSMANN, *Compt. rend.*, 236 (1953) 2118.
13 D. H. MARRIAN, V. L. SPICER, M. E. BALIS AND G. B. BROWN, *J. Biol. Chem.*, 189 (1951) 533.
14 A. MARSHAK, *J. Biol. Chem.*, 189 (1951) 607.
15 W. M. MCINDOE AND J. N. DAVIDSON, *Brit. J. Cancer*, 6 (1952) 200.
16 J. N. DAVIDSON AND R. M. S. SMELLIE, *Biochem. J.*, 52 (1952) 594.
17 J. N. DAVIDSON AND R. M. S. SMELLIE, *Biochem. J.*, 52 (1952) 599.
18 G. W. CROSBIE, R. M. S. SMELLIE AND J. N. DAVIDSON, *Biochem. J.*, 54 (1953) 287.
19 H. A. DELUCA, R. J. ROSSITER AND K. P. STRICKLAND, *Biochem. J.*, 55 (1953) 193.
20 P. S. OLMSTED AND C. A. VILLEE, *J. Biol. Chem.*, 212 (1955) 179.
21 A. MARSHAK AND H. J. VOGEL, *J. Biol. Chem.*, 189 (1951) 597.

[22] W. S. VINCENT, *Proc. Nat. Acad. Sci.*, 38 (1952) 139.
[23] R. THOMAS, *Biochim. Biophys. Acta*, 8 (1952) 71.
[24] R. MARKHAM AND J. D. SMITH, *Biochem. J.*, 46 (1950) 513.
[25] R. MARKHAM AND J. D. SMITH, *Biochem. J.*, 49 (1951) 401.
[26] C. A. KNIGHT, *J. Biol. Chem.*, 197 (1952) 241.
[27] F. L. BLACK AND C. A. KNIGHT, *J. Biol. Chem.*, 202 (1953) 51.
[28] R. W. DORNER AND C. A. KNIGHT, *J. Biol. Chem.*, 205 (1953) 959.
[29] W. D. COOPER AND H. S. LORING, *J. Biol. Chem.*, 211 (1954) 505.
[30] G. HAMOIR, *Biochem. J.*, 50 (1951) 140.
[31] Y. KHOUVINE AND M. L. HIRSCH, *Bull. soc. chim. biol.*, 34 (1952) 305.
[32] Y. KHOUVINE AND H. DE ROBICHON-SZULMAISTER, *Bull. soc. chim. biol.*, 34 (1952) 1050, 1056.
[33] C. H. PARSONS, JR., *Arch. Biochem. Biophys.*, 47 (1953) 76.
[34] L. EGGMAN, S. J. SINGER AND S. G. WILDMAN, *J. Biol. Chem.*, 205 (1953) 969.
[35] J. D. WATSON AND F. H. C. CRICK, *Nature*, 171 (1953) 737.
[36] M. H. F. WILKINS, A. R. STOKES AND H. R. WILSON, *Nature*, 171 (1953) 738.
[37] R. E. FRANKLIN AND R. G. GOSLING, *Nature*, 171 (1953) 740.
[38] M. H. F. WILKINS, W. E. SEEDS, A. R. STOKES AND H. R. WILSON, *Nature*, 172 (1953) 759.
[39] E. CHARGAFF, *Experientia*, 6 (1950) 201.
[40] C. F. CRAMPTON, R. LIPSHITZ AND E. CHARGAFF, *J. Biol. Chem.*, 211 (1954) 125.
[41] E. CHARGAFF, in Fibrous Proteins and Their Biological Significance, *Symposia Soc. Exp. Biol.*, 9 (in the press).
[42] G. SCHMIDT, in E. CHARGAFF AND J. N. DAVIDSON, *The Nucleic Acids*, New York, 1 (1955) 555.
[43] L. PAULING, R. B. COREY AND H. R. BRANSON, *Proc. Nat. Acad. Sci.*, 37 (1951) 205.
[44] W. T. ASTBURY AND F. O. BELL, *Nature*, 141 (1938) 747.
[45] N. W. PIRIE, *Biochem. J.*, 56 (1954) 83.
[46] E. CHARGAFF AND E. VISCHER, *Ann. Rev. Biochem.*, 17 (1948) 201.

Received February 12th, 1955