Characterization of nucleotide arrangement in deoxyribonucleic acids through stepwise acid degradation*

We present here a procedure capable of uniquely characterizing and contrasting DNA specimens that cannot be distinguished otherwise. It is based on a standardized and quantitative method for the estimation of the 3',5'-diphosphates of cytidine (pCp) and thymidine (pTp) produced at three defined steps of acid degradation of DNA. That some pPy.p is formed during partial acid hydrolysis of DNA has been known since the early work of LEVENE[1] and THANNHAUSER[2] and was confirmed recently[3].

The types of nucleotide sequence that could serve as precursors of pPy.p were studied first in model experiments with several dinucleotides[4]. As expected ApCp and CpAp gave rise, readily and quantitatively, to pCp and Cp, respectively (5 min, 100°, 0.01 *M* H_2SO_4); effects that may be extended to the trinucleotide sequence Pu.pPy.pPu.p. Increase of the acid concentration to 0.2 *M* led to an appreciable production of pCp from the dinucleotide CpCp, 33% of the doubly esterified C appearing as pCp in 120 minutes. Under the same conditions, the dinucleotide mixture CpTp and TpCp (containing 70% of the first) gave rise to 49% of the potential pTp and to 12% of the potential pCp.

These experiments together with kinetic studies on the rates of hydrolysis of the pyrimidine nucleosides and their phosphates seem to indicate that the production of pTp or pCp or both from pyrimidine dinucleotides follows the same pattern as the breakdown of purine–pyrimidine dinucleotides, namely, the initial fission of one of the two glycosidic bonds, with C more labile than T. A secondary, slower formation of diphosphates can, therefore, be expected from pyrimidine dinucleotide precursors originating during the acid degradation of DNA.

In extending the study of the production of pTp and pCp to DNA preparations from different cellular sources[4] conditions had to be found permitting the approximate, but standardizable, differentiation between pPy.p arising from trinucleotide segments, in which pyrimidine is flanked by purines, and pPy.p yielded by pyrimidine oligonucleotide fragments among which dinucleotides will be the principal contributors. Model experiments with dinucleotides led to the choice of three stages of hydrolysis for the differential analysis of the distribution density of pyrimidine nucleotides in different DNA preparations: I. 30 minutes; II. 60 minutes; III. 120 minutes; all in 0.1 *M* H_2SO_4, 100°. The rapid production of diphosphates in stage I essentially reflects the number of Pu.pPy.pPu.p sequences; their subsequent slower formation, as shown by the increments during the following stages, is indicative of the type of predominant all-pyrimidine precursors.

About 10 mg of DNA are required for a complete analysis; 0.5 ml of acid was used for hydrolysis. Portions of the solution removed at the completion of each stage were placed on Dowex-2-formate columns (200–400 mesh, 11 × 0.6 cm) and fractionally eluted with formic acid and ammonium formate[5]. Identification of separated components was made by spectrophotometry and paper chromatography in *iso*butyric acid–ammonium *iso*butyrate (pH 3.6)[6].

We have so far examined ten different DNA specimens, both fractions and total preparations, of which three have been chosen for inclusion in Table I, since they are among those that cannot be distinguished by total constituent analysis. All figures are based on at least duplicate hydrolyses, with the values rarely deviating by more than 5% from the reported averages.

An inspection of the data provided in the table will show that the differential analysis procedure described here is a useful tool for the characterization of the nucleotide arrangement in DNA preparations, even if they are indistinguishable by other means. Taking into account our entire survey, which will later form the subject of a detailed publication, the following range of figures may be quoted for stage I that can be taken as an approximation of the relative proportions of purine-flanked pyrimidine nucleotides in different DNA preparations. The quantity of cytidine diphosphate varied from 6.0 to 15.5% of the total cytidylic acid; the corresponding figures for thymidine diphosphate were 12.8 and 23.5%.

The most obvious general feature of the results is the nonidentity of the values of T/C and pTp/pCp, contrary to what would have been expected from a randomly constructed polynucleotide. A more direct indication of nonrandomness emerges when the experimental values for DNA pyrimidines appearing as diphosphates at 30 minutes (varying from 9.8 to 17.3% of total pyrimidines) are contrasted with the figure of 25 ± 1.6% for the percentage of single pyrimidine nucleotides flanked by purine nucleotides obtained by the mathematical treatment of random distribution of nucleotides in a linear sequence[8].

---

* Abbreviations: DNA, deoxypentose nucleic acids; Pu., Py., A, G, C, T, deoxyribosides of purines, pyrimidines, adenine, guanine, cytosine, thymine, respectively. The position of the esterified phosphoric acid is indicated by p being placed at the right of the symbol if it is attached at 5' of the nucleoside, at the left if it is on 3'.

TABLE I

DIFFERENTIAL ANALYSIS OF PYRIMIDINE DISTRIBUTION IN DNA*

	Starting material								
*DNA fraction***	*Ox (0.75)*			*Man (1.0)*			*Arbacia lixula (2.6)*		
A, G, C, T; as mole % P	31.1; 19.0; 20.0; 29.9			30.9; 19.4; 19.8; 29.9			31.0; 18.5; 19.6; 30.9		
T/C, molar ratio	1.50			1.51			1.58		
	Stepwise hydrolysis								
*Stage****	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*
pTp, as mole % Tp	15.9	19.5	22.6	19.9	21.1	24.3	16.0	20.2	24.2
pCp, as mole % Cp	9.2	11.7	13.2	6.3	9.8	12.0	14.0	17.1	17.7
pTp/pCp, molar ratio	2.58	2.49	2.56	4.77	3.26	3.07	1.80	1.86	2.15

* The mole % figures refer to moles per 100 moles of indicated constituent present in DNA.
** The figures in parentheses describe the NaCl molarity at which the particular fraction was isolated through the fractional dissociation of histone nucleate[7].
*** The values for nucleoside diphosphates liberated during the hydrolysis stages I, II, III (30, 60, 120 minutes, respectively, in 0.1 *M* H_2SO_4 at 100°) have been corrected for the decomposition of the diphosphates occurring in the course of this treatment.

Other fragments, separated and in some cases estimated, include pCpCp, pCpTp, pTpTp, pCpCpTp, and pTpTpCp.

Our results permit the conclusion that the detailed aspects of the arrangement of pyrimidine nucleotides (and therefore of purine nucleotides) vary widely in DNA specimens of different origin, providing a new means of distinction; that in all cases examined, however, the distribution of constituents appears far from random; and that—not unlike a previous suggestion[9]—about 70% of the DNA pyrimidines occur as oligonucleotide tracts containing three or more pyrimidine nucleotides in a row.

This work will form part of a dissertation to be submitted by H. S. SHAPIRO in partial fulfilment of the requirements for the degree of Doctor of Philosophy, Columbia University. It was supported by research grants from the Rockefeller Foundation, the American Cancer Society and the U.S. Public Health Service, to which latter agency one of us (H.S.S.) is indebted for a Predoctorate Fellowship.

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

HERMAN S. SHAPIRO
ERWIN CHARGAFF

1 P. A. LEVENE, *J. Biol. Chem.*, 48 (1921) 119.
2 S. J. THANNHAUSER AND B. OTTENSTEIN, *Z. physiol. Chem.*, 114 (1921) 39.
3 C. A. DEKKER, A. M. MICHELSON AND A. R. TODD, *J. Chem. Soc.* (1953) 947.
4 H. S. SHAPIRO AND E. CHARGAFF, *Federation Proc.*, 15 (1956) 352.
5 W. E. COHN, in E. CHARGAFF AND J. N. DAVIDSON, *The Nucleic Acids*, Vol. I, Academic Press, Inc., New York, 1955, p. 211.
6 B. MAGASANIK, E. VISCHER, R. DONIGER, D. ELSON AND E. CHARGAFF, *J. Biol. Chem.*, 186 (1950) 37.
7 C. F. CRAMPTON, R. LIPSHITZ AND E. CHARGAFF, *J. Biol. Chem.*, 211 (1954) 125.
8 A. M. MOOD, *Ann. Math. Stat.*, 11 (1940) 367.
9 C. TAMM, H. S. SHAPIRO, R. LIPSHITZ AND E. CHARGAFF, *J. Biol. Chem.*, 203 (1953) 673.

Received December 3rd, 1956