Reprinted from

Biochimica et Biophysica Acta, 312 (1973) 633–636

BBA 97730

PYRIMIDINE TRACTS OF THE (A+T)-RICH SATELLITE DNA FROM *CANCER PAGURUS*

STANISLAV D. EHRLICH*, JEAN-PAUL THIERY and GIORGIO BERNARDI

Laboratoire de Génétique Moléculaire, Institut de Biologie Moléculaire, Faculté des Sciences, Paris 5° (France)

(Received February 23rd, 1973)

SUMMARY

The pyrimidine tracts of (A+T)-rich (3 % G+C) satellite DNA from *Cancer pagurus* have been analyzed. Base pairs involving isolated and clustered T (essentially T_2 and T_3) form close to 89 and 7 % of the DNA, respectively. The latter are responsible, together with G · C base pairs (about 65 % of which are in isolated positions) for the incomplete renaturability of the satellite DNA. The presence of non-alternating dA · dT sequences explains the anomalously low density of the satellite.

INTRODUCTION

In spite of the wide interest raised by the (A+T)-rich satellite DNAs of *Crustacea* since their discovery by Sueoka[1], direct information about their nucleotide sequences is limited to the fact that in *Cancer borealis* 93 % of the satellite DNA is in the form of alternating dA–dT sequences, as judged by nearest neighbor analysis[2]. Here we report analytical data on the pyrimidine tracts of the (A+T)-rich satellite DNA from *Cancer pagurus* showing that: (a) non-alternating dA · dT (essentially $(dA \cdot dT)_2$ and $(dA \cdot dT)_3$) forms about 7 % of the DNA; (b) about 65 % of G · C base pairs are in isolated positions; (c) the presence of the non-alternating dA · dT explains the "anomalous" buoyant density of the satellite and is responsible, together with the G · C base pairs, for its incomplete renaturability.

MATERIALS AND METHODS

DNA from the testes of *C. pagurus* (animals were obtained in Roscoff, France) was prepared by the detergent procedure[3]. The (A+T)-rich satellite was prepared by equilibrium centrifugation in Cs_2SO_4–Hg^{2+} density gradients[4]. Synthetic poly-[d(A–T) · d(A–T)] was obtained from Miles (Elkhart, Ind.). Buoyant density measure-

* Present address: Department of Genetics, Lt. Joseph P. Kennedy Jr Laboratories for Molecular Medicine Stanford University, Medical School, Stanford, Calif. 94305, U.S.A.

ments in CsCl density gradients were performed as already described[5] using two density markers at the same time, phage 2C DNA ($\rho = 1.742$ g/cm^3) and *Escherichia coli* DNA ($\rho = 1.710$ g/cm^3). The G+C contents of the satellite DNA was determined by nucleoside analysis after enzymatic degradation[6], and found to be equal to 3 %. Depurination and isostich analysis were performed as described earlier[8] (Method A).

RESULTS AND DISCUSSION

Four isostich classes, corresponding to mono, di, tri and pentanucleotides, were separated by chromatography on DEAE-cellulose columns (Fig. 1) of depurinated (A+T)-rich satellite DNA. The first three were digested to nucleosides and analyzed as described previously[8]; a typical separation of T and C is shown in Fig. 2. Table I shows the relative amounts and the compositions of the isostich classes.

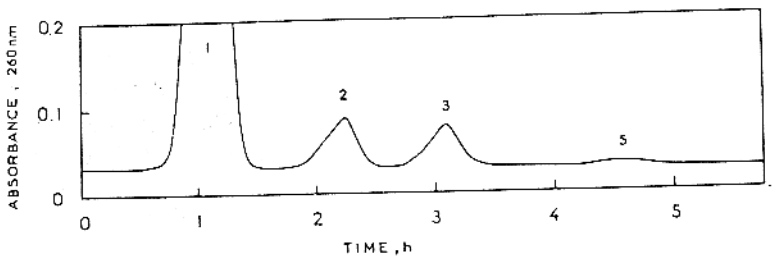

Fig. 1. Chromatography of a formic acid–diphenylamine hydrolysate of the (A+T)-rich satellite DNA (5 $A_{270\,nm}$ units of pyrimidines) on a 0.4 cm × 6 cm DEAE-cellulose column. Elution was performed with a linear gradient of ammonium acetate buffer pH 5.3 (0.05 to 0.675 M; total vol. 80 ml). Numbers refer to isostich classes.

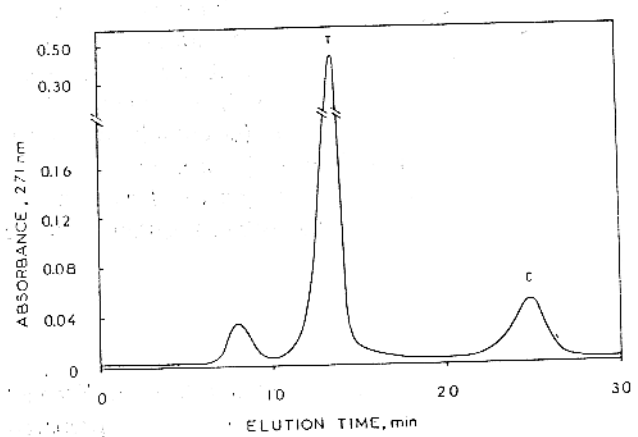

Fig. 2. Bio-Gel P2 chromatography of an isostich class digested simultaneously with spleen exonuclease and phosphatase. 0.03 $A_{270\,nm}$ units of pyrimidine nucleosides were separated on a 0.5 cm × 8 cm column. Elution was performed with 2 mM ammonium carbonate buffer, pH 10.4, using a flow rate of 6 ml/h. Absorbance at 271 nm was quantitatively recorded; relative quantities of T and C were calculated from surfaces under the corresponding peaks. The small peak preceding T is composed of the ultraviolet absorbing material contaminating the isostich fractions.

89 % of the satellite DNA is present in alternating [d(A–T) · d(A–T)] sequences; about 7 % appears to be in non-alternating sequences 2 or 3 nucleotides long. The relative amounts of C increase from mono- to di- and trinucleotides. Pentamers were also present, but they were too scarce to be analyzed, their frequency being of the order of one pentamer per 1000 nucleotides. The complete absence of tetranucleotides confirms the view that the pentamers originate from the satellite DNA and not from a hypothetical contamination by nuclear DNA. These results differ from previous data[9] indicating that *C. pagurus* satellite DNA contains exclusively alternating dA–dT sequences, a conclusion drawn on the basis of less sensitive analytical methods, but confirm the suggestion[10], based on the formation of thymine dimers upon ultraviolet irradiation, that as much as 5 % of the satellite DNA might be in the form of non-alternating dA · dT base pairs.

TABLE I

DISTRIBUTION AND CYTIDINE CONTENTS OF PYRIMIDINE ISOSTICHS IN CRAB SATELLITE DNA

Pyrimidine isostich	(C+T)* %	C** %
1	91.2	2 (2.2)
2	4.2	0.24 (5.6)
3	3.5	0.31 (8.8)
4	0	—
5	0.4	n.d.
6 and longer	0	—

* (C+T) and C are given as percentages of total pyrimidines in the DNA.
** Values in parentheses give the percentages of C in the isostichs.

The buoyant densities of *C. pagurus* (A+T)-rich satellite DNA and synthetic poly[d(A–T) · d(A–T)] were found to be equal to 1.679 and 1.680 g/cm^3, respectively. It appears, therefore, that the contribution of G · C base pairs is counterbalanced by that of the very light non-alternating dA · dT sequences. Knowing the contribution[11] of G · C base pairs and of alternating poly[d(A–T) · d(A–T)] to the density, and assuming that they are additive, one can calculate that the contribution of the nonalternating dA · dT dimers and trimers, interestingly, is identical to that of poly-(dA · dT), 1.647 g/cm^3 (ref. 12). A similar situation has already been described for yeast mitochondrial DNAs[13,14].

The incomplete renaturability of the (A+T)-rich satellite[4,9,10,14] is more understandable, as already suggested[10], in terms of the presence of both non-alternating dA · dT sequences and G · C base pairs (representing all together 10 % of the satellite DNA) rather than on the basis of the latter alone, since a larger percentage of the satellite is bound to be mismatched after cooling or neutralization.

It is very likely that non-alternating dA · dT sequences also exist in the (A+T)-rich satellite DNAs of other crustaceans .This is suggested from the nearest-neighbor data of Schwarz *et al.*[2] for the satellite of *C. borealis*; in addition, the circular dichroism data of Allen *et al.*[15] indicate the presence of 5 % A · A and T · T in the satellite of *C. antennarius*.

REFERENCES

1 Sueoka, N. (1961) *J. Mol. Biol.* 3, 31–40
2 Swartz, M. N., Trautner, T. A. and Kornberg, A. (1962) *J. Biol. Chem.* 237, 1961–1967
3 Kay, E. R. M., Simmons, N. S. and Dounce, A. L. (1952) *J. Am. Chem. Soc.* 74, 1724–1726
4 Davidson, N., Widholm, J., Nandi, U. S., Jensen, R., Olivera, B. M. and Wang, J. C. (1965) *Proc. Natl. Acad. Sci. U.S.* 53, 111–118
5 Bernardi, G., Piperno, G. and Fonty, G. (1972) *J. Mol. Biol.* 65, 173–189
6 Carrara, M. and Bernardi, G. (1968) *Biochim. Biophys. Acta* 155, 1–7
7 Piperno, G. and Bernardi, G. (1971) *Biochim. Biophys. Acta* 238, 388–396
8 Ehrlich, S. D., Thiery, J. P. and Bernardi, G. (1972) *J. Mol. Biol.* 65, 207–212
9 Pochon, F., Massoulié, J. and Michelson, A. M. (1966) *Biochim. Biophys. Acta* 119, 249–257
10 Skinner, D. M. (1967) *Proc. Natl. Acad. Sci. U.S.* 58, 103–110
11 Schildkraut, C. L., Marmur, J. and Doty, P. (1962) *J. Mol. Biol.* 4, 430–443
12 Szybalski, W. (1968) in *Methods in Enzymology*, (Grossman, L. and Moldave, K., eds), Vol. 12, part B, pp. 330–360
13 Bernardi, G., Carnevali, F., Nicolaieff, A., Piperno, G. and Tecce, G. (1968) *J. Mol. Biol.* 37, 493–505
14 Bernardi, G., Faures, M., Piperno, G. and Slonimski, P. (1970) *J. Mol. Biol.* 48, 23–42
15 Allen, F. S., Gray, D. M., Roberts, G. P. and Tinoco, Jr, I. (1972) *Biopolymers* 11, 853–879