REPETITIVE PRIMATE DNA CONTAINING THE RECOGNITION SEQUENCES FOR TWO RESTRICTION ENDONUCLEASES WHICH GENERATE COHESIVE ENDS

Peter GRUSS and Gerhard SAUER

Institut für Virusforschung, Deutsches Krebsforschungszentrum, 69 Heidelberg, West Germany

Received 26 September 1975

1. Introduction

Physical maps of the genomes of small viruses have been constructed using restriction endonucleases [1,2] and the more complex DNA of higher organisms has been studied recently. Digestion of bovine, human and murine DNA followed by electrophoretic separation of the cleavage products revealed, besides diffusely distributed unique sequences, fragments of defined length [3–7].

When DNA from CV-1 cells, a permanent line derived from *Cercopithecus aethiops* monkey kidney cells, was treated with various restriction endonucleases, it turned out that a repetitive component of this DNA was susceptible to attack by both Eco R_I and Hind III enzymes. Thus, we were able to map the distance between the respective recognition sites within a monomeric sequence. Furthermore, partial digests permitted an estimation of the size of the randomly arranged repeated DNA sequences.

A more practical aspect of this work concerns the repeated DNA fragments that can be obtained after double digestion with both Eco R_I and Hind III endonucleases. Such fragments bear cohesive ends [8–11] and can serve as useful 'linkers' in molecular engineering experiments where DNA sequences of different origins bearing either the Eco R_I or the Hind III termini are to be joined with each other.

2. Materials and methods

CV-1 DNA and SV40 DNA were isolated and purified as described [12]. Digestion of the DNA with Eco R_I enzyme was performed in 12 mM $MgCl_2$ and 90 mM Tris, pH 7.9, and with Hind III enzyme in 6 mM mercaptoethanol, 6 mM Tris and 50 mM NaCl, pH 7.4, at 37°C. The reactions were terminated by addition of one seventh the reaction volume of a solution containing 8 M urea, 0.1 M EDTA, and 0.02% bromophenol blue in 50% sucrose. The samples were electrophoresed in 1.4% agarose gels, or, in polyacrylamide gels (4% acrylamide, 0.2% bisacrylamide) as described [13].

3. Results and discussion

CV-1 DNA was reacted with restriction nucleases for various periods of time, and then electrophoresed together with a Hind III digest of SV40 DNA (fig.1,b) as marker in an agarose gel. After electrophoresis, the gel was stained with ethidium bromide and photographed under ultraviolet light. Both Eco R_I (fig.1,a) and Hind III nucleases (fig.1,c and d) generate, besides larger diffusely distributed CV-1 DNA fragments in the upper region of the gel, a number of discrete bands, which appear as integral multiples of one monomer DNA fragment. The size of the fragments (indicated in fig.1) can be estimated using Hind III restricted SV40 DNA (fig.1,b) as a reference [1]. The monomer consists of 170 ± 10 base pairs, regardless of whether Eco R_I or Hind III nuclease had been employed. On the other hand, Hind II nuclease as well as Hemophilus parainfluenzae nuclease I and II do not cleave CV-1 DNA into fragments of defined length.

Regarding the DNA sequences that contain the recognition sites for both Eco R_I and Hind III nuclease, it appeared from the work of Maio [14] that component αDNA which is repeated over

a million times per cell could be the possible candidate. Isolation of rapidly reassociating component αDNA as described [14] and treatment of this fraction with both Eco R_I and Hind III nucleases revealed, that this DNA does indeed harbor the recognition sites for both enzymes. In agreement with Maio's findings, we could estimate from densitometer tracings of the gels that about 15 to 20% of the *Cercopithecus aethiops* genome is comprised by component αDNA.

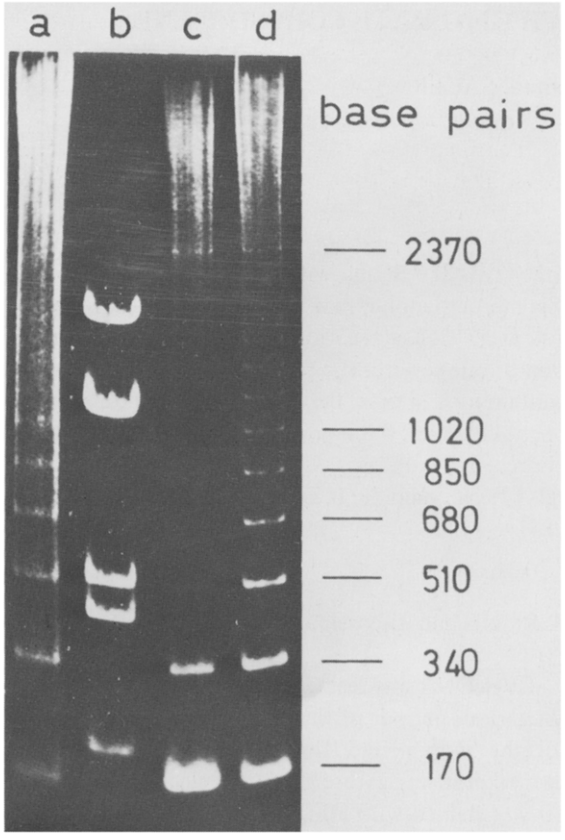

Fig.1. Agarose gel electrophoresis of CV-1 DNA and SV40 DNA fragments produced by Eco R_I and Hind III restriction endonucleases. (a) CV-1 DNA (6.8 μg) was digested with 5 μl of Eco R_I for 20 h; (b) SV40 DNA (5.2 μg) isolated from strain Va 45–54 was digested with 5 μl of Hind III for 2 h at 37°C; (c) CV-1 DNA (6.8 μg) was mixed with 2×10^4 cpm (0.02 μg) [^{32}P] SV40 DNA (strain Va 45–54) and digested with 5 μl of Hind III for 20 h; (d) reaction mixture as described under (c), digestion stopped after 3 h. The volumes of the reaction mixtures were 70 μl each. After electrophoresis the gel was stained for 5 min in electrophoresis buffer containing 5 μg/ml ethidium bromide. Photographs were taken using ultraviolet light for fluorescence. The actual distance between the origin at the top and the smallest fragment is 18 cm. The size of the fragments (expressed in base pairs) was estimated (assuming 650 daltons per base pair) from a plot relating log molecular weight to relative distance migrated using Hind III restricted SV40 DNA (mol. wt 3.2×10^6 daltons) as a reference [1].

The arrangement of repetitive DNA sequences within mammalian DNA is of considerable interest. Partial digests of CV-1 DNA (such as the one shown in fig.1,d) have revealed tandemly repeated sequences as large as 5000 to 5500 base pairs long. However, owing to the large amounts of superimposed unique DNA in the upper portions of the gels this may be an underestimation, and the tandem repeats might be even larger.

Treatment with Hind III nuclease leads to the appearance of a discrete band of another DNA class consisting of approximately 2370 base pairs (fig.1,c and d). Eco R_I nuclease fails to cleave the DNA at these sites, as this band is missing in the gel (fig.1,a). Possibly, components β or γ [15] may be represented within this class of DNA.

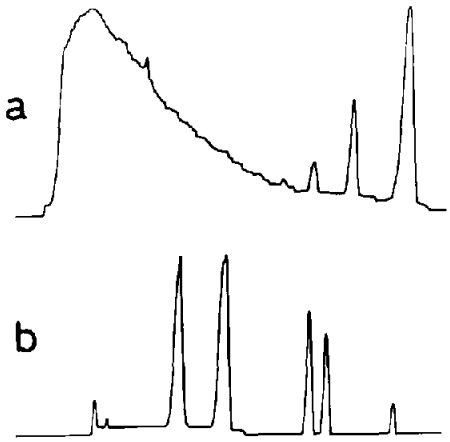

Fig.2. Densitometer tracings of Hind III restricted CV-1 and SV40 DNA. (a) Densitometer tracing of the photograph of Hind III digested CV-1 DNA shown in fig.1c; (b) densitometer tracing (Joyce Loebl Chromoscan) of an autoradiograph of Hind III digested [^{32}P] SV40 DNA which had been added as internal marker to the gel containing the CV-1 DNA shown in fig.1c. Agfa-Gevaert film Curix RPI was exposed for 3 days and then processed. The different peaks in the tracing represent (from left to right): linear SV40 DNA (Form III), superhelical SV40 DNA (Form I) (both comprising 3.4% of the total DNA), Hind III fragment A, Hind III fragments B + C (which are not resolved here), Hind III fragments D, E and F.

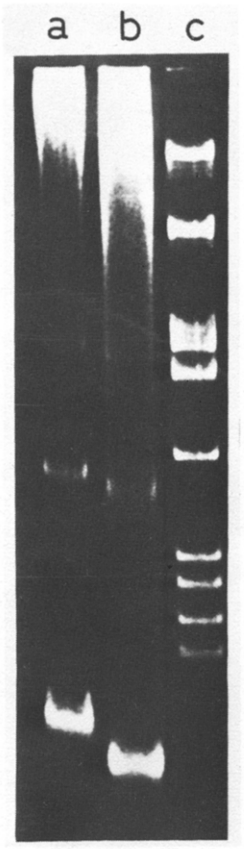

Fig.3. Sequential digestion of CV-1 DNA with Hind III and Eco R_I nucleases. Photographs of an ethidium bromide stained polyacrylamide gel. CV-1 DNA (25 μg) was reacted with 20 μl of Hind III for 20 h at 37°C. An aliquot (7μg) was reacted subsequently for 6 h at 37°C with 10 μl of Eco R_I. (a) CV-1 DNA digested with Hind III; (b) CV-1 DNA digested with Hind III and with Eco R_I; (c) SV40 DNA (4 μg) strain 777 digested for 20 h with Hind III + Hind II (10 μl each).

Attempts have been made to monitor the degree of digestion by including, prior to addition of restriction nucleases, labelled SV40 DNA in the reaction mixture and evaluating the extent to which SV40 DNA was digested. In the digest shown in fig.1,c, almost 100% of the ^{32}P-labelled SV40 DNA had been cleaved by Hind III. This is evidenced in fig.2 which shows the densitometer tracings of the restricted CV-1 DNA (fig.2,a) and the distribution of the digested ^{32}P-SV40 DNA (fig.2b) within the same gel. Despite almost complete digestion of SV40 DNA (fig.2,b) about 22 ± 2% of the repetitive DNA persisted as dimers and about 5 ± 2% as trimers that could not be converted by Hind III nuclease to monomers. Dimers and trimers were isolated by homogenization of the gel and re-digested with Hind III nuclease. No further conversion to monomers was achieved. Because of the above-mentioned reasons we conclude, that these are repetitive sequences containing mutated Hind III recognition sites.

The presence of both the Eco R_I and the Hind III recognition sequences within component αDNA permits, by double digestion, the localization of both sites relative to each other. A polyacrylamide gel electrophoresis pattern of Hind III digested CV-1 DNA (fig.3,a), Hind III + Eco R_I digested CV-1 DNA (Fig.3,b) and Hind II + Hind III cleaved SV40 DNA as a marker (Fig.3,c) revealed that the Eco R_I cleavage site is localized about 20 ± 10 base pairs in from the end of the Hind III generated monomer DNA sequence. This conclusion is derived from the acceleration of the double-digested monomer relative to the distance migrated by the Hind III monomer.

It can also be seen that the mutated dimers and trimers must also be affected in their internal Eco R_I sites since they are not further converted by Eco R_I to monomers. Rather, one of their distal Eco R_I recognition sites (one only per dimer, or, trimer, respectively) is attacked, thus rendering the dimers and trimers smaller by 20 base pairs each.

Since the oncogenic SV40 replicates in CV-1 cells and since the viral genome also integrates into the CV-1 DNA the data reported in this paper permit a variety of experiments concerning the recombination between viral and mammalian DNA. For example, one can test now whether or not the SV40 genome integrates into the highly repetitive component αDNA (manuscript in preparation). One can also study whether the substituted SV40 DNA molecules [16] which presumably arise by an integration-excision event and which contain covalently bound host cell DNA also harbour component αDNA.

Furthermore, the non-coding repetitive DNA fragments obtained after double digestion with Eco R_I and Hind III enzymes can be employed as linkers in genetic engineering experiments to perform joinings between DNA sequences bearing either the Eco R_I or the Hind III generated cohesive termini.

Acknowledgements

We thank Drs Mayer and Goebel, Gesellschaft für Molekularbiologische Forschung, Stöckheim, and Institut für Mikrobiologie der Universität Würzburg for a generous gift of Eco R_I enzyme.

References

[1] Danna, K. J., Sack, G. H. and Nathans, D. (1973) J. Mol. Biol. 78, 363–367.
[2] Griffin, B. E., Fried, M. and Cowie, A. (1974) Proc. Natl. Acad. Sci. USA 71, 2077–2081.
[3] Philippsen, P., Streeck, R. E. and Zachau, H. G. (1974) Eur. J. Biochem. 45, 479–488.
[4] Roizès, G. (1974) Nucleic Acids Research 1, 1099–1120.
[5] Hörz, W., Hess, I. and Zachau, H. G. (1974) Eur. J. Biochem. 45, 501–512.
[6] Southern, E. M. (1975) J. Mol. Biol. 94, 51–69.
[7] Botchan, M. R. (1974) Nature 251, 288–292.
[8] Mertz, J. E. and Davis, R. W. (1972) Proc. Natl. Acad. Sci. USA 69, 3370–3374.
[9] Hedgpeth, J., Goodman, H. M. and Boyer, H. W. (1972) Proc. Natl. Acad. Sci. USA 69, 3448–3452.
[10] Roy, P. and Smith, H. O. (1973) J. Mol. Biol. 81, 445–459.
[11] Murray, K. and Old, R. W. (1974) in: Progr. Nucl. Acid. Res., Academic Press, London 14, 117–185.
[12] Waldeck, W., Kammer, K. and Sauer, G. (1973) Virology 54, 452–464.
[13] Chowdhury, K., Gruss, P., Waldeck, W. and Sauer, G. (1975) Biochem. Biophys. Res. Commun. 64, 709–716.
[14] Maio, J. J. (1971) J. Mol. Biol. 56, 579–595.
[15] Kurnit, D. M. and Maio, J. J. (1973) Chromosoma 42, 23–36.
[16] Lavi, S. and Winocour, E. (1972) J. Virol. 9, 309–316.