NUCLEOTIDE SEQUENCE AROUND THE REPLICATION ORIGIN OF POLYOMA VIRUS DNA

Eiichi SOEDA, Kin-ichiro MIURA, Akiko NAKASO* and Genki KIMURA+

*National Institute of Genetics, Mishima 411, *Department of Virology, Tottori University School of Medicine, Yonago 683* and *+Cancer Institute, Kyushu University School of Medicine, Fukuoka 812, Japan*

Received 23 May 1977

1. Introduction

The origin of the DNA replication of polyoma virus has been localized recently at the 71±3 map unit numbered from the cleavage site with *Eco*RI endonuclease [1] by two independent methods [2,3]. The DNA replication starts from this point and proceeds bidirectionally at a similar rate on the circular DNA [2,3]. Also both the 5′-termini of 'early' and 'late' viral mRNAs are located at regions close to the origin of DNA replication [4]. Therefore, the region around the replication origin holds the key for the correlation between the initial steps in DNA replication and RNA synthesis.

*Hpa*II endonuclease cleaves polyoma DNA into fragments, and their relative positions have been mapped [2]. Among these, fragment 5 (*Hpa*II-5) includes the region in question, since it extends from 70.8–78.5 map units. This fragment is cleaved once at the 72.8 map unit by *Hha*I endonuclease. The initiation site for DNA replication is probably on or close to the resulting shorter fragment [5] (fig.1B).

We deduced the nucleotide sequence of both strands of this short fragment by recently developed sequencing methods using enzyme and chemical reagents. The fragment is composed of 73 base pairs and the sequenced strands were perfectly complementary to each other when arranged in an antiparallel manner. A symmetrical A–T-rich stretch was flanked by a G–C-rich region with a 2-fold rotational axis of symmetry. This structural feature is also observed in the fragment (*Eco*RII-G) of SV40 DNA containing the origin of replication [6].

Mailing address: Dr K. Miura, National Institute of Genetics, Mishima 411, Japan

2. Experiments and discussion

2.1. *Preparation of DNA and its restriction fragments*

Since most of the relaxed DNA molecules separated by CsCL–ethidium bromide centrifugation of Hirt's extract [7] were found to be viral DNA, this DNA isolation method was used for the DNA preparation without further purification steps. *Hap*II endonuclease has the same recognition site as that of the *Hpa*II enzyme [8] and cuts the polyoma DNA into eight discrete fragments (fig.1A), giving the same cleavage pattern as the DNA from the A3 strain [2,3]. The strain LP147-2 originates from the large plaque of Vogt and Dulbecco [9] and was passaged 4 times in primary baby mouse kidney cell cultures consecutively at low multiplicity of infection in our studies. These fragments were designated as *Hap*II-1–8 in order of decreasing size, corresponding to *Hpa*II fragments 1–8, respectively. The *Hap*II-5 fragment was cut once at the 72.8 map unit with *Hha*I endonuclease into two fragments which were separated from each other on a 8% polyacrylamide gel. The resulting shorter fragment *Hap*II-5/*Hha*I-1, which is located close to the *Hap*II-3-5 junction, is most probably the fragment containing the replication origin, and therefore, it was sequenced.

The late strand of *Hap*II-5/*Hha*I-1, complementary to late mRNA, was selectively labelled at the 5′-terminus as follows. Whole digestion products of polyoma DNA with *Hap*II endonuclease were treated with bacterial alkaline phosphatase to delete terminal phosphate,

Exposed overnight

6 days

3% gel

8% gel

Fig.1A

Fig.1. Cleavage pattern with restriction enzymes and genetic map of the region around the origin of DNA replication. (A) The polyoma virus, LP147-2 strain, which originated from the large plaque strain, was provided by Vogt and Dulbecco [9] and purified further through diluted passages of the virus. Exponentially growing 3T6 mouse cells were infected with the polyoma virus at 1 p. f. u./cell in Eagle's medium modified by Vogt and Dulbecco supplemented with 2% calf serum. At 96 h after infection with 1–3 medium changes during incubation period, viral DNA was selectively extracted from the infected cells [7] and purified by a CsCl–ethidium bromide centrifugation. For labelling with ^{32}P, the infected cultures were washed once with the phosphate-free medium at 24 h after infection and further incubated in the medium supplemented with dialyzed 2% calf serum containing carrier-free [^{32}P]orthophosphate. Supercoiled and relaxed viral DNA were cut with *Hap*II and *Hha*I endonucleases in R buffer (30 mM Tris–HCl, pH 7.5, containing 7 mM $MgCl_2$ and 7 mM β-mercaptoethanol and separated by electrophoresis in a 3–8% polyacrylamide slab gel (0.2 × 20 × 40 cm) in E buffer [10]. The active bands were visualized by autoradiography after drying the gel. The upperpart of the photograph was obtained after overnight exposure to X-ray film and the lower after 6 day's exposure. Supercoiled and relaxed viral DNA were digested with *Hap*II (a,c) and with a mixture of *Hap*II and *Hha*I (b,d).

and then these DNA fragments were labelled with ^{32}P at the 5′-termini using T_4-induced polynucleotide kinase [11] and [γ-^{32}P]ATP [12]. The labelled fragments were digested by *Hha*I endonuclease and the digests separated by electrophoresis in an 8% polyacrylamide gel. The smaller fragment *Hap*II-5/*Hha*I-1 should carry ^{32}P at the 5′-terminus of only one strand, the late strand, whereas the larger fragment *Hap*II-5/*Hha*I-2 should carry ^{32}P at the 5′-terminus of the other strand, the early strand. Here the former fragment (*Hap*II-5/*Hha*I-1) was used for late strand analysis.

If DNA is first digested by *Hha*I and labelled with ^{32}P at the 5′-terminus, the following digestion by *Hap*II yields DNA fragments, which are labelled in the complementary strand, the opposite of the above experiment. Thus the *Hha*I-1/*Hap*II-5 fragment should carry ^{32}P at the 5′-terminus of the early strand. Therefore, analysis of *Hap*II-5/*Hha*I-1 and *Hha*I-1/*Hap*II-5 provides information on the nucleotide sequence of either the late strand or the early strand of the same DNA fragment.

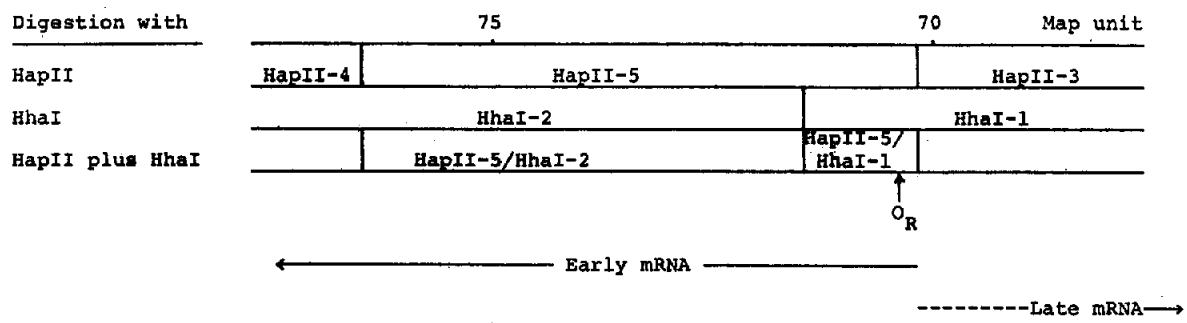

Fig.1(B). The recognition site of *Hap*II endonuclease is the same as *Hpa*II endonuclease and cuts polyoma DNA into eight discrete fragments. *Hha*I endonuclease cuts the DNA into four fragments. These fragments are designated *Hap*II-1–8 and *Hha*I-1–4 in order of decreasing size. The origin of viral DNA replication (OR) is in *Hap*II-5 proximal to the *Hap*II-3–5 junction [1,2]. *Hap*II-5 is cut once with *Hha*I on the side of this junction and therefore the resulting shorter fragment *Hap*II-5/*Hha*-1 is most probably the fragment containing the replication origin. Both 5′-termini of early and late mRNA are located in the proximity of the origin [4].

2.2. *Sequencing with the use of chemical reagents*

For sequencing the DNA fragments, we used the dimethylsulfate–hydrazine technique developed by Maxam and Gilbert [13]. The procedure depends on base specific but partial cleavage of labelled fragments and separation of the resulting fragments on electrophoresis in a denaturation gel of polyacrylamide.

Figures 2a and 2b show the cleavage patterns of *Hap*II-5/*Hha*I-1. Nucleotide residues are numbered from the *Hap*II cleavage site of the late strand which starts with 5′-CGG– of the recognition site of *Hap*II, 5′-CCGG-3′[8]. In the autoradiogram (fig.2a) the terminal C does not appear because it ran off the gel. The intensity of bands in the 'A' lane is a little faint. This may be due to poor reactivity of adenosine residues in double-stranded DNA towards methylation because the N4 position of adenosine residues is blocked by a hydrogen bond and the *N*-glucoside bond is not so sensitive as that in guanosine residues. Even so, the nucleotide sequence from position 2–53 could be deduced from the autoradiogram.

*Hha*I recognizes the sequence of 5′-GCGC-3′ producing the fragment carrying the last C as the new 5′-terminus [14], which is labelled later with polynucleotide kinase. The labelled early strand of *Hha*I-1/*Hap*II-5 was isolated from the whole digestion products labllled at the 5′-terminus followed by a second digestion with *Hap*II (fig.2c). The bands in the 'A' lane become clear with slight modifications including longer exposure to acid after methylation of denatured fragments. Several oligonucleotides near the 5′-termini were missing, judging from the position of the dye marker where decanucleotides were stuated. The nucleotide sequence of the early strand from position 13–65 was determined from this photograph.

2.3. *Sequencing with the use of DNAase I*

The nucleotide sequence from the cleavage site of *Hha*I was confirmed by two dimensional finger printing of partial digestion products of labelled *Hha*I-1/*Hap*II-5 fragments with DNAaseI (fig.3). The nucleotide sequence from position 59–71 was deduced

Fig.2. Sequencing with the use of the dimethylsulfate–hydrazine technique. The *Hap*II-5/*Hha*I-1 fragment labelled at only one 5′-terminus (1–2 × 10^4 cpm Cerenkov radiation) was treated at 20°C for 45 min with dimethylsulfate (purine specific cleavage) or with hydrazine in the absence (pyrimidine specific cleavage) and the presence of 1 M NaCl (C-specific cleavage). Preferential cleavage at the A residue was performed by exposure of the methylated fragment to 0.1 M HCl. The modified fragment was specifically cleaved by a piperidine catalyzed β-elimination reaction at 90°C for 30 min. The cleaved products were analyzed on a denaturation gel of 20% polyacrylamide. Electrophoresis was carried out at 600 V for 14 h (b and c) and for 24 h (a). The wet gel was wrapped in a plastic sheet and exposed to X-ray film at −80°C for 10–14 days. The numbers beside the nucleotides start from the *Hap*II cleavage site at the *Hap*II-3–5 junction. BPB and XC show the reference positions of bromphenol blue and xylene cyanol XX markers.

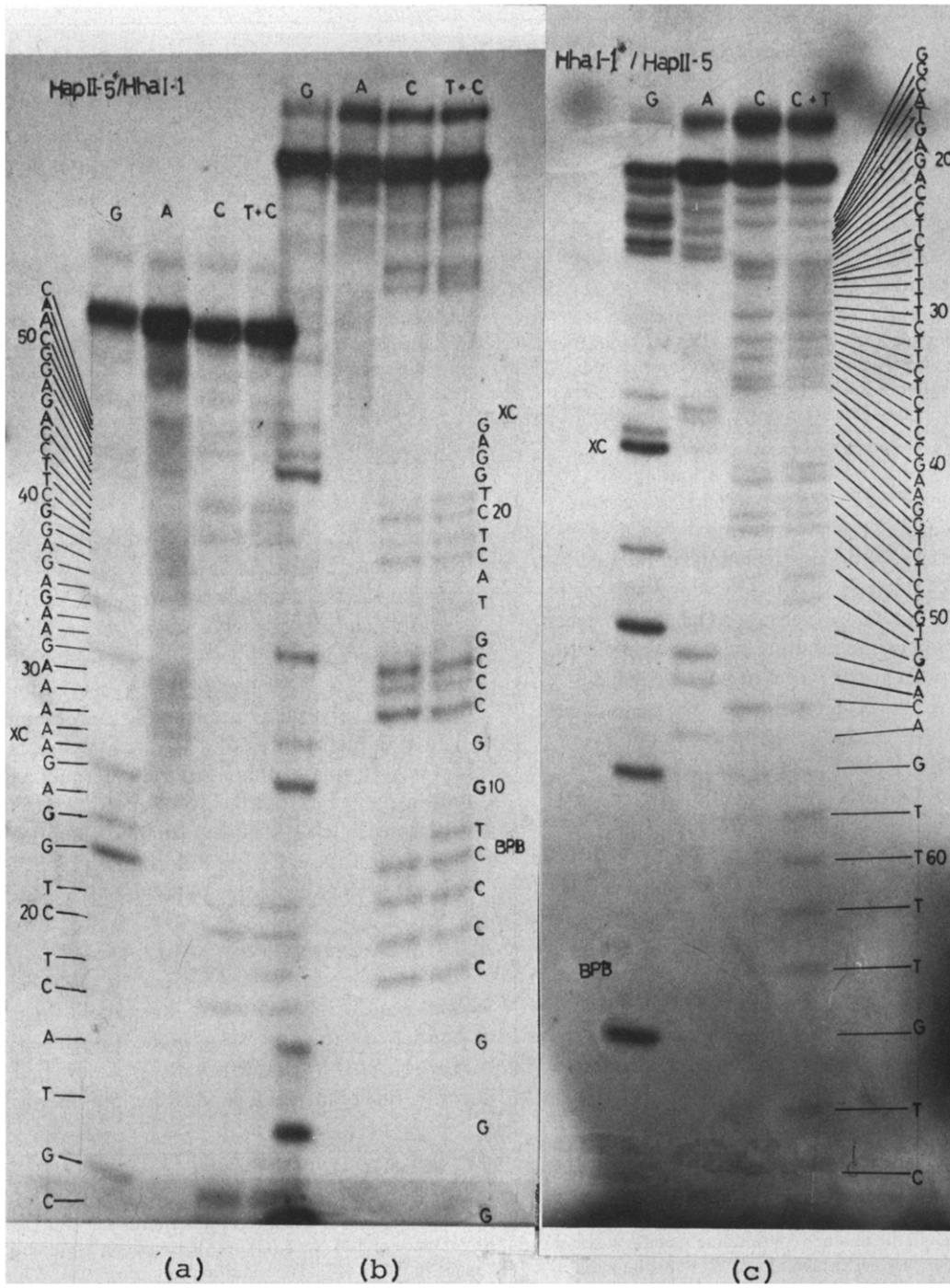

Fig.2

from the mobility shifts [17] on the fingerprint. Of these, the arrangement of CTGTTTT overlapped with that shown in fig.2c (position 59–65). The eight nucleotide sequence from the *Hap*II cleavage site was also determined from the mobility shifts, overlapping with GGGCCC shown in fig.2a, although the data are not shown here.

2.4. *Construction of the nucleotide sequence and possible transient folded structure*

It is possible to construct the entire nucleotide sequence of fragment *Hap*II-5/*Hha*I-1 by a combination of chemical and enzymatic partial digestion methods (fig.4A). When each strand sequenced was arranged in an antiparallel manner, the base sequences in both

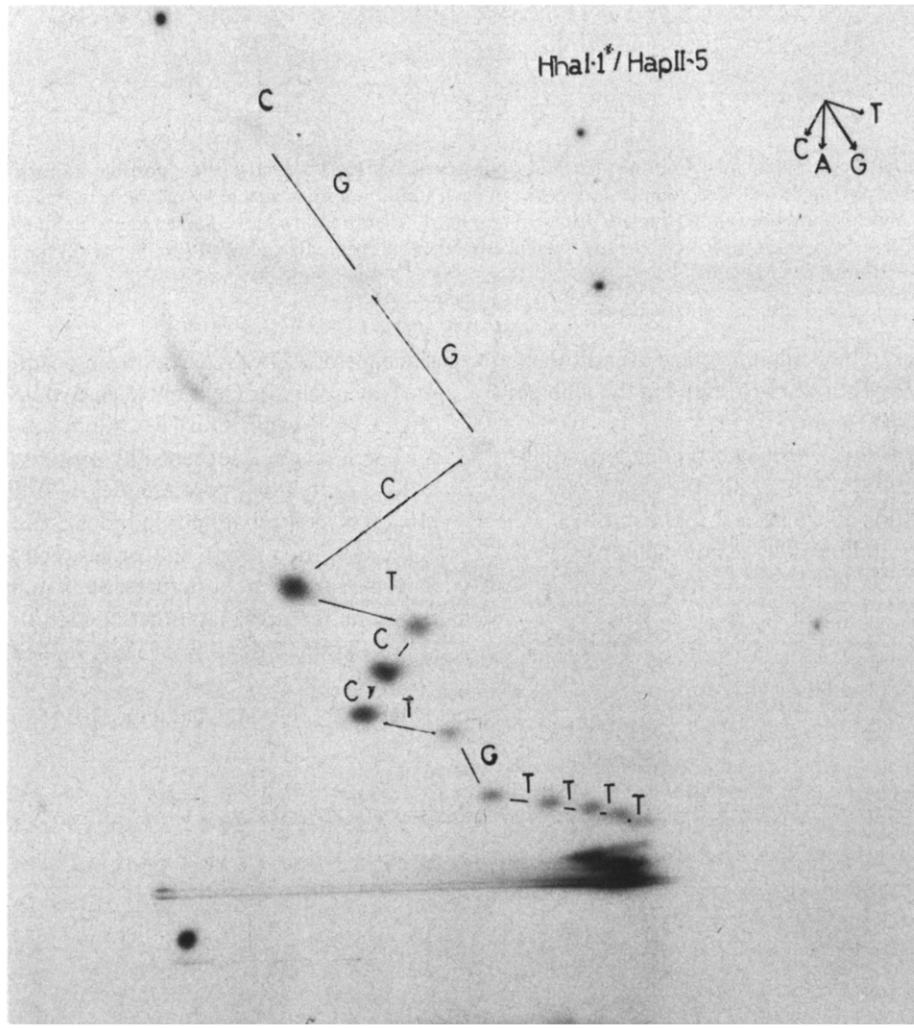

Fig.3. Sequencing according to mobility shifts of partial digestion products with DNAase. The fragment labelled at the 5′-terminus of the early strand, *Hha*I-1*/*Hap*II-5 was digested with 3.3 ng DNAase I (Worthington electrophoretically purified) in 6 μl of R buffer containing 1 μg T_7 DNA. Aliquots of 1 μl were withdrawn from the reaction mixture at the times of 10 min, 20 min, 40 min and 60 min and frozen immediately. The frozen aliquots were combined and fractionated by electrophoresis on cellulose acetate, at pH 3.5, followed by homochromatography on a PEI plate as described by Volckaert et al. [15].

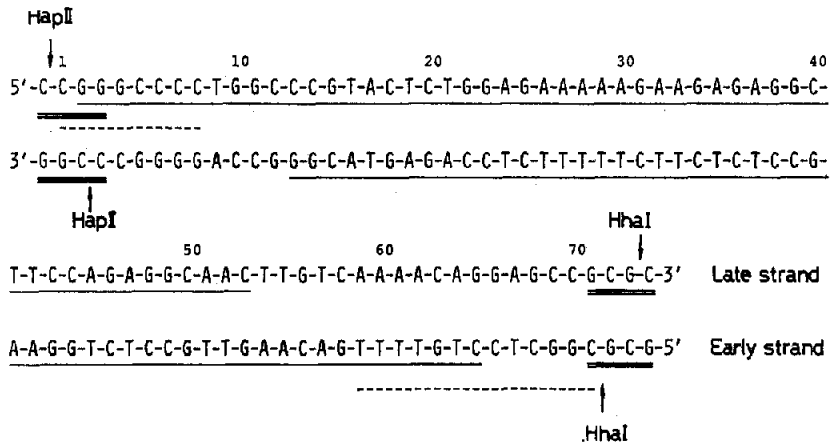

Fig.4. Construction of the complete nucleotide sequence of both strands of *Hap*II-5/*Hha*I-1 and a possible transient structure containing double hairpin loops. (A). The complete sequence of each strand was constructed by overlapping two sequences obtained from analysis of chemical (solid line) and enzymatic (dotted line) partial digestion products. Some oligonucleotides were added at both termini of the fragment, taking into account the sequences recognized by *Hap*II and *Hha*I (double line). The nucleotides were numbered from the cleavage site of *Hap*II.

strands were seen to be complementary over all the DNA fragment *Hap*II-5/*Hha*I-1, resolving the ambiguities caused by faint bands.

The sequenced fragment may be characterized by several features as shown in fig.4B; the G–C-rich region from position 1–15 has a 2-fold rotational axis of symmetry, followed by a true palindromic sequence of AGAAAAAGA from position 24–32. The A–T-rich palindrome is between two inverted repeats of CTCTGGA (position 18–24 and 42–48). If the A–T-rich region is loosened by proteins like a non-histone protein with preferential A–T-binding capacity [16], it is likely that hairpin loops could be formed in the G–C-rich region and/or between the inverted repeats as transient structures. Such remarkable structural features may function effectively in specific events like the initiation of DNA replication and RNA synthesis.

Recently Subramanian et al. [6] reported the

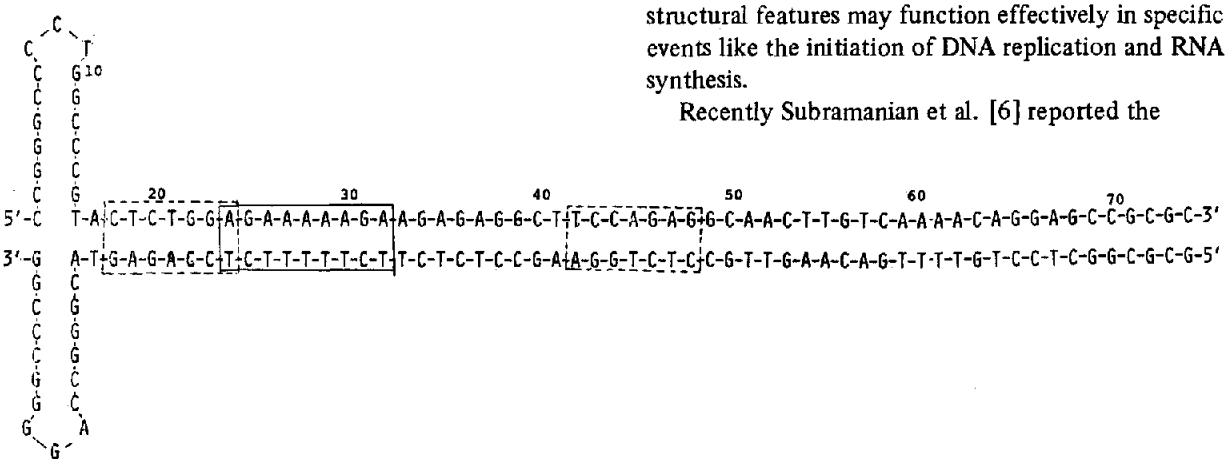

Fig.4(B). The G–C-rich stretch extending from position 1–15 has a 2-fold rotational axis of symmetry and therefore can form transient double hairpin structures. The A–T-rich stretch with the true palindromic sequences is shown in a box with solid lines which is between two inverted repeats (boxes with dotted lines). The early strand is presumed to be transcribed from right to left starting from this fragment or proximal to this fragment.

complete nucleotide sequence of the *Eco*RII G fragment of SV40 DNA, which contains the origin of DNA replication and also specifies the 5′-termini of early and late viral mRNAs. The fragment contains characteristic symmetric sequences as well as the inverted repeated sequences. They have assigned two regions as signal sites involved in the initiation of DNA replication and transcription. It is significant to compare the nucleotide sequence near the replication origin of DNA between SV40 and polyoma virus, because both the viruses can infect common mammalian cells like mouse embryo cells and produce viral early mRNA. Although complete homology in base sequence was not detected, there is a common feature, namely that the A-T-rich true palindromic sequence is flanked by G-C-rich hairpin loops, corresponding to the proposed promoter site of early mRNA of SV40 DNA.

Acknowledgements

We thank Dr Mitsuru Takanami for kindly providing restriction enzymes, *Hha*I, *Hap*II and advice on the fragmentation of DNA with the enzymes. We also gratefully acknowledge Dr Takeo Maruyama for programming of the computer to seek characteristic structures like the palindromic sequence.

References

[1] Folk, W. R. and Wang, H.-C. E. (1974) Virology 61, 140–155.
[2] Griffin, B. E., Fried, M. and Cowei, A. (1974) Proc. Natl. Acad. Sci. USA 71, 2077–2081.
[3] Crawford, L. V., Robbins A. K., Nicklin, P. M. and Osborn, K. (1974) Cold Spring Harb. Symp. Quant. Biol. 39, 219–225.
[4] Kamen, R. and Shure, H. (1976) Cell 7, 361–371.
[5] Miller, L. K. and Fried, M. (1976) J. Virol. 18, 824–832.
[6] Subramanian, K. N., Dhar, R. and Weissman, S. M. (1977) J. Biol. Chem. 252, 355–367.
[7] Hirt, B. (1967) J. Mol. Biol. 26, 365–369.
[8] Takanami, M., Sugimoto, K., Sugisaki, H. and Okamoto, T. (1976) Nature 260, 297–302.
[9] Vogt, M. and Dulbecco, R. (1960) Proc. Natl. Acad. Sci. USA 46, 365–370.
[10] Danna, K. J., Sack, G. H. and Nathans, D. (1973) J. Mol. Biol. 78, 363–376.
[11] Panet, A., Van de Sande J. H., Loewen, P. C., Khorana, H. G., Raae, A. J., Lillehaug, J. R. and Kleppe, K. (1973) Biochemistry 12, 5045–5050.
[12] Schendel, P. F. and Wells, R. D. (1973) J. Biol. Chem. 248, 8319–8321.
[13] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[14] Roberts, R. J., Myers, P. A., Morrison, A. and Murray, K. (1976) J. Mol. Biol. 103, 199–208.
[15] Volckaert, G., Min jou, W. and Fiers, W. (1976) Anal. Biochem. 72, 433–446.
[16] Thomas, L. T. and Pater, G. L. (1976) Proc. Natl. Acad. Sci. USA 73, 4364–4368.
[17] Maniatis, T., Jeffrey, A. and Kleid, D. G. (1975) Proc. Natl. Acad. Sci. USA. 72, 1184–1188.