# The EcoRI Site of Simian Virus 40 Deoxyribonucleic Acid: Nucleotide Sequences of the Minus and Plus Strands

David E. Garfin, Jo-Ann C. Leong*, and Howard M. Goodman*
Departments of Microbiology and Biochemistry & Biophysics*
University of California Medical Center
San Francisco, California 94143

Received November 18,1975

SUMMARY: The nucleotide sequences of the minus and plus strands of simian virus 40 DNA in a 17-base-pair segment spanning the EcoRI restriction site have been identified. The minus-strand sequence is (5') T-G-G-C-G-A-G-A-A-T-T-C-C-T-T-T-G and the plus-strand sequence is its complement: (5') C-A-A-A-G-G-A-A-T-T-C-T-C-G-C-C-A.

Simian virus 40 DNA (SV40 DNA) contains a single, unique site for the EcoRI restriction endonuclease (1, 2, 3) which is used as a reference position in the mapping of the viral genome (4, 5). The site is located in the segment of DNA transcribed late in lytic infection (6, 7) and is 0.33 fractional lengths clockwise from the origin of DNA replication (8, 9). DNA complementation analyses with temperature-sensitive mutants of SV40 (10) indicate that the EcoRI site is within the gene defined by either the B or C complementation group (11). Although the product of this gene has not been identified, temperature-sensitive mutants of the B and C types are defective in virion protein formation or assembly at restrictive temperatures.

The substrate for the EcoRI endonuclease is the symmetrical, double-stranded equivalent of the hexamer G-A-A-T-T-C (12, 13). Since the twofold symmetry of the substrate does not extend to the other sequenced base pairs, the primary structures of the two SV40 DNA strands at the EcoRI site can be distinguised from one another. We have determined which of the two EcoRI-site sequences is that of the minus strand and which is the plus-strand sequence. The minus strand of SV40 DNA is complementary to the asymmetric RNA synthesized in vitro by Escherichia coli RNA polymerase from supercoiled SV40 DNA templates (14) and is

the strand transcribed into stable RNA during the early phase of the lytic cycle of SV40 infection (15, 16, 17). We demonstrate here that the sequence T-G-G-C-G-A-G-A-A-T-T-C-C-T-T-T-G is the sequence on the minus strand of SV40 DNA which overlaps the EcoRI site and that the plus-strand sequence is its complement.

## METHODS

The preparation of supercoiled SV40 DNA, its cleavage with EcoRI endonuclease and subsequent 5'-terminal labeling are described elsewhere (13, 18). EcoRI endonuclease (3), free of contaminating nucleolytic activities, was provided by R. B. Meagher. Assymetric transcription of supercoiled SV40 DNA by E. coli DNA-dependent RNA polymerase has been described (14). The assymetry of the cRNA synthesized was confirmed by self-reassociation and RNase digestion of the RNA as described by Lindstrom and Dulbecco (15). To separate the strands of SV40 DNA, 5 μg of [5'-$^{32}$P] labeled, EcoRI-cleaved SV40 DNA (about $10^6$ cpm) were first boiled for 10 minutes in 1 ml $H_2O$, then cooled to 85° C, and 100 μg of cRNA were added. After 5 minutes at 85°C, the solution was cooled to 68°C and adjusted to 0.05M NaCl. Hybridization at 68°C was for 1 hour ($C_o t = 7.25 \times 10^{-4}$ mole-sec/liter, $C_r t = 1.89 \times 10^{-2}$ mole-sec/liter). Following dilution with water to 5 ml, the nucleic acids were adsorbed to hydroxyapatite (19). Single-stranded plus-strand DNA was eluted with 0.15 M sodium phosphate buffer, pH 6.8, and the double-stranded complex of minus-strand DNA hybridized to cRNA was eluted with 0.4 M sodium phosphate buffer, both at 60°C. The RNAs in the eluates were hydrolyzed in 0.6 N NaOH for 3 hours at 37°C, after which the solutions were neutralized and adjusted to 1.6M NaCl. The separated DNAs were self-annealed in the resultant solutions for 16 hours at 68°C ($C_o t$ = 0.9 mole-sec/liter), diluted into 500 ml $H_2O$, adsorbed to hydroxyapatite, and eluted as before. The single-stranded material from this procedure (the 0.15 M-sodium-phosphate eluates) was purifed further by a second round of self-annealing. All 0.4 M-sodium-phosphate eluates were combined as a double-stranded control. The three samples--minus and plus strands (2 ml each) and control (4 ml)--were individually dialyzed against 0.01M triethylamine bicarbonate, pH 8, 0.0001 M EDTA. After four changes of 1.5 liters each, the dialysates were air dried, then dried again three times from water. The three DNA samples were each digested to small, random-sized oligonucleotides by incubation with pancreatic DNase, desalted on small columns of DEAE-Sephadex A25 (bicarbonate form) and dried. The DNAse digestion products were fingerprinted in two dimensions by electrophoresis on cellulose acetate and homochromatography using conditions detailed in the legend to Fig. 1.

## RESULTS AND DISCUSSION

Supercoiled SV40 DNA was cleaved with EcoRI restriction endonuclease and the resultant full-length, linear molecules were labeled with $^{32}$P-phosphate groups at the 5'-termini of the two DNA strands. This labeled DNA was then denatured and the minus strand was hybridized to a 20-fold mass excess of complementary RNA (cRNA). The DNA-cRNA hybrid was separated from plus-strand DNA by batchwise adsorption to and elution from hydroxyapatite,

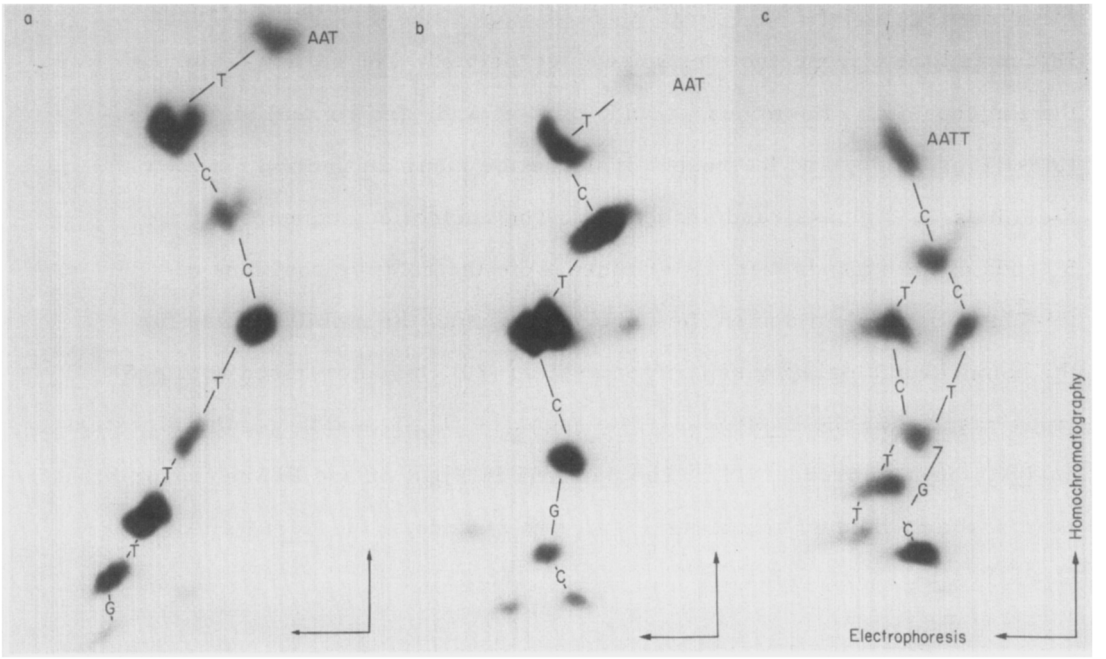


Fig. 1. Pancreatic DNase fingerprints of the separated strands of SV40 DNA at the EcoRI site. Intact SV40 DNA was digested to completion with EcoRI restriction endonuclease in 0.1 M Tris-HCl, pH 7.5, 0.005 M $MgCl_2$, 0.05 M NaCl. 5'-terminal phosphates were removed with bacterial alkaline phosphatase and replaced with $^{32}$P-phosphates from [$\gamma$-$^{32}$P]rATP (300 mCi/$\mu$mole) in the polynucleotide kinase reaction. The strands of 5 $\mu$g of the labeled, linear DNA were separated as described in the Methods, purified free of complementary sequences by two cycles of self-annealing, and the double-stranded material from the two self-annealing steps was combined and used as a control. The DNAs were each digested with 100 $\mu$l of pancreatic DNase at 0.1 mg/ml in 0.01 M Tris-HCl, pH 7.4, 0.02 M $MgCl_2$ for 40 minutes at 37°C and desalted on small columns of DEAE-Sephadex A25 (bicarbonate form). The samples were separated in two dimensions, first by high-voltage electrophoresis on cellulose acetate in 7 M urea, 10% acetic acid, 0.001 M $Na_2$-EDTA, pH 3.5, then by homochromatography on 1:7.5 DEAE-cellulose: cellulose thin layers developed with homomixture c (7 M urea, 0.001 M EDTA, containing 3% yeast RNA previously hydrolyzed for 30 minutes with 1 N KOH) (13, 23). The radioautograph of the fractionation of the DNase digest of purified minus-strand DNA is shown in panel a, the pattern for the plus strand is in panel b, and the control pattern is in panel c. All labeled oligonucleotides have 5'-phosphates. The symbols for all phosphates have been omitted. The letters represent the 3'-terminal residues by which two adjacent oligonucleotides, connected by lines, differ.

---

and cRNA was removed from the minus-strand DNA by alkaline hydrolysis. Single-stranded minus- and plus-strand DNAs were further purified by twice self-annealing each strand preparation and removing any resultant double-stranded DNA by hydroxyapatite chromatography. The purified single-

stranded DNAs were digested to small oligonucleotides with pancreatic DNase and the oligonucleotides were fingerprinted. The salient feature of the fingerprints of the minus-strand, plus-strand, and control oligonucleotides (Fig. la, lb, and lc, respectively) is the clear distinction between the minus- and plus-strand sequences. The nucleotide sequences at the 5'-ends of the strands were determined from the relative positions of [5'-$^{32}$P] oligonucleotides on the fingerprints using the mobility rules for the homochromatography separation system (20), by comparison with previous results for the EcoRI substrate sequence (12,13) and by partial digestion analysis of each product (13). The patterns in Figs. la and lb are characteristic of the 5'-labeled sequences on the two strands at the EcoRI site of SV40 DNA as shown below.

```
                    ↓
5'  -T-G-G-C-G-A-G-A-A-T-T-C-C-T-T-T-G-  3'  (minus)
3'  -A-C-C-G-C-T-C-T-T-A-A-G-G-A-A-A-C-  5'  (plus)
                          ↑
```

The arrows indicate the internucleotide bonds cleaved by the EcoRI endonuclease. The sequences are written with the minus strand reading 5'-to-3' clockwise, in accordance with the conventional mapping scheme for SV40 DNA (7, 8). Fig. lc, the fingerprint of the control sample, contains both sequences. A few oligonucleotides from the minus strand contaminate the fingerprint of plus-strand DNA (Fig. lb). Since the plus strand is obtained in hydroxyapatite chromatography as the non-hybridized or single-stranded DNA species, and since prolonged storage of 5'-labeled DNA led to an increase in the intensity of the non-specific spots, it is possible that these oligonucleotides are produced from minus-strand DNA fragments which do not hybridize under our conditions and elute as single-stranded molecules. Nevertheless, the patterns are clearly distinct enough to allow unambiguous strand assignment.

It is not known whether the region adjacent to the EcoRI site has any definite coding function. The *in vitro* translation studies of Roberts *et al.* (21) indicate that the EcoRI site of SV40 DNA is not part of the cistron of the

major capsid protein since coat-protein production can be programmed by EcoRI-cleaved SV40 DNA. The site seems to be in the "gene" defined by temperature-sensitive mutations of the B or C type (11). B and C mutants synthesize SV40 DNA but fail to produce complete virions at non-permissive temperatures. In an analysis in which fragments obtained from wild-type SV40 DNA were used to complement the DNA of several temperature-sensitive mutants, Lai and Nathans (10) have shown that B and C mutations map on both sides of the EcoRI site. Moreover, Carbon et al. (22) have constructed SV40 DNA molecules with deletions spanning the EcoRI site and these can be complemented by temperature-sensitive A mutants but not by B mutants. The segment of DNA described here has the capacity to code for five or six amino acids depending on the reading frame. However, until it is definitely shown that the site is actually part of a structural gene, any assignment of codons is merely hypothetical.

ACKNOWLEDGMENTS: This work was supported by U. S. Public Health Service grants AI 00299 and CA 14026. J. C. L. is the recipient of an American Cancer Society, California Division, Dernham Postdoctoral Fellowship.

## REFERENCES

1. Mulder, C. and Delius, H. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 3215-3219.
2. Morrow, J. F. and Berg, P. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 3365-3369.
3. Greene, P. J., Betlach, M. C., Goodman, H. M. and Boyer, H. W. In R. B. Wickner (ed.) Methods in molecular biology, vol. 7: DNA replication. pp. 87-111, Marcel Dekker, Inc. New York.
4. Danna, K. J., Sack, G. H. Jr., and Nathans, D. (1973) J. Mol. Biol. 78, 363-376.
5. Subramanian, K. N., Pan, J., Zain, S. and Weissman, S. M. (1974) Nucleic Acids Research 1, 727-752.
6. Khoury, G., Martin, M. A., Lee, T.N.H., Danna, K. J. and Nathans, D. J. Mol. Biol. 78, 377-389.
7. Sambrook, J., Sugden, B., Keller, W. and Sharp, P. A. (1973) Proc. Nat. Acad. Sci. U.S.A. 70, 3711-3715.
8. Danna, K. J. and Nathans, D. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 3097-3100.
9. Fareed, G. C., Garon, C. F. and Salzman, N. P. (1972) J. Virol. 10, 484-491.
10. Lai, C. J., and Nathans, D. (1974) Cold Spring Harbor Symp. Quant. Biol. 39, 63-60.
11. Martin, R. G., Chou, J. Y., Avila, J. and Saral, R. (1974) Cold Spring Harbor Symp. Quant. Biol. 39, 17-24.

12. Hedgpeth, J., Goodman, H. M. and Boyer, H. W. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 3448-3452.
13. Garfin, D. E., Boyer, H. W., and Goodman, H. M. (1975) Nucleic Acids Research 2, 1851-1965.
14. Westphal, H. (1970) J. Mol. Biol. 50, 407-420.
15. Lindstrom, D. M. and Dulbecco, R. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 1517-1520.
16. Sambrook, J., Sharp, P. A. and Keller, W. (1972) J. Mol. Biol. 70, 57-71.
17. Khoury, G. and Martin, M. A. (1972) Proc. Nat. Acad. Sci. U.S.A. 69, 1925-1928.
18. Dugaiczyk, A., Hedgpeth, J., Boyer, H. W. and Goodman, H. M. (1974) Biochemistry 13, 503-512.
19. Leong, J., Garapin, A., Jackson, N., Fanshier, L., Levinson, W., and Bishop, J. M. (1972) J. Virol. 9, 891-902.
20. Sanger, F., Donelson, J. E., Coulson, A. R., Kossel, H. and Fischer, D. (1973) Proc. Nat. Acad. Sci. U.S.A. 70, 1209-1213.
21. Robert, B. E., Gorecki, M., Mulligan, R. C., Danna, K. J., Rozenblatt, S. and Rich, A. (1975) Proc. Nat. Acad. Sci. U.S.A. 72, 1922-1926.
22. Carbon, J., Shenk, T. E. and Berg, P. (1975) Proc. Nat. Acad. Sci. U.S.A. 72, 1392-1396.
23. Brownlee, G. G. and Sanger, F. (1969) Eur. J. Biochem. 11, 395-399.