JOURNAL OF VIROLOGY, June 1979, p. 936-941
0022-538X/79/06-0936/06$02.00/0

Vol. 30, No. 3

Nucleotide Sequence Analysis of Two Simian Virus 40 Mutants with Deletions in the Region Coding for the Carboxyl Terminus of the T Antigen

HUGO VAN HEUVERSWYN,[1] CHARLES COLE,[2] PAUL BERG,[3] AND WALTER FIERS[1]*

Laboratory of Molecular Biology, State University of Ghent, B-9000 Ghent, Belgium[1]; Department of Human Genetics, School of Medicine, Yale University, New Haven, Connecticut 06510[2]; and Department of Biochemistry, Stanford University Medical Center, Stanford, California, 94305[3]

Received for publication 28 November 1978

Nucleotide sequence analysis of two simian virus 40 early mutants, *dl*1263 and *dl*1265, which lack a DNA segment around map positions 0.21 and 0.18, respectively (C. Cole, T. Landers, S. Goff, S. Manteuil-Brutlag, and P. Berg, J. Virol. 24:277-294, 1977), revealed in-phase deletions of 33 nucleotide pairs for *dl*1263 and 39 nucleotide pairs for *dl*1265. The 33-base-pair deletion in *dl*1263 does not correspond to an apparent shortening by 6,000 daltons observed for the mutant T antigen. In *dl*1265 the normal termination signal as well as most of the proline-rich terminal tryptic peptide has been removed, and the carboxyl terminus of the mutant T antigen is a series of three cysteine residues.

The simian virus 40 (SV40) DNA early region is at least 2,574 base pairs long and extends counterclockwise from approximately 0.65 to 0.17 map units. The early region codes for two proteins, small-t and large-T antigen (11, 14, 15), which are initiated at the same AUG start codon corresponding to map position 0.649 (12). The corresponding ATG sequence on the DNA fixes a reading frame that is open up to the information for a termination codon at position 0.547, and a polypeptide with molecular weight of 20,503, which presumably corresponds to small-t, can be deduced from the DNA nucleotide sequence (19). Large-T has a molecular weight of 94,000, and the large-T gene comprises two noncontiguous segments of the early region: the first extends from 0.649 to about 0.59 and the second segment runs from about 0.535 to 0.174 (1, 5, 6, 12, 18, 19). At least two additional virus-induced early antigens, U antigen and tumor-specific transplantation antigen, have been detected both in lytic infections and in transformed cells; they are coded for by the distal and middle portion, respectively, of the SV40 early region (ref. 20 and references therein). Moreover, SV40 expresses a function which helps human adenovirus to grow in monkey cells. By isolation and characterization of adenovirus-SV40 hybrids, it was shown that this helper function corresponds to the carboxyl-terminal part of the large-T antigen (8, 17), of which the coding DNA sequence was recently published (16, 18).

Two viable early-region mutants, *dl*1263 and *dl*1265, lack a DNA segment in this C-terminal region. The position of the deletion in *dl*1263 was estimated around 0.21 map units (the exact map position of this deletion is in fact 0.20), and its large-T antigen is apparently 6,000 daltons lighter than wild-type large-T; the deletion in *dl*1265 maps around position 0.18 and apparently does not affect the size of large-T antigen (4), although in the absence of alkylation the latter seems to be more prone to aggregation effects (C. Cole, L. V. Crawford, and P. Berg, submitted for publication). Both *dl*1263 and *dl*1265 are derived from the viable mutant *dl*861 (4).

To characterize these deletions at the nucleotide level, each mutant DNA was digested separately with restriction endonucleases *Hin*dII + III, *Hae*III, and *Alu*I (Fig. 1 and 2). *dl*1263 gave rise to a shorter (relative to wild type) *Hin*dII + III fragment B and a shorter *Alu*I fragment A. The characteristic bands for *Hae*III fragments C and D were missing in the restriction cleavage pattern. The resulting new band was not detected on the gel, but it presumably comigrated with the *Hae*III fragment B. We conclude that *dl*1263 lacks the *Hae*III cleavage site at map position 0.195.

Upon digestion with the respective enzymes, *dl*1265 DNA also resulted in a shorter *Hin*dII + III fragment B and in a shorter *Alu*I fragment A. Furthermore, digestion with *Hae*III enzyme revealed a shorter fragment C, thus localizing the *dl*1265 deletion between the *Hin*dII + III G/B junction (map position 0.169) and the *Hae*III cleavage site at map position 0.195 (Fig. 2).

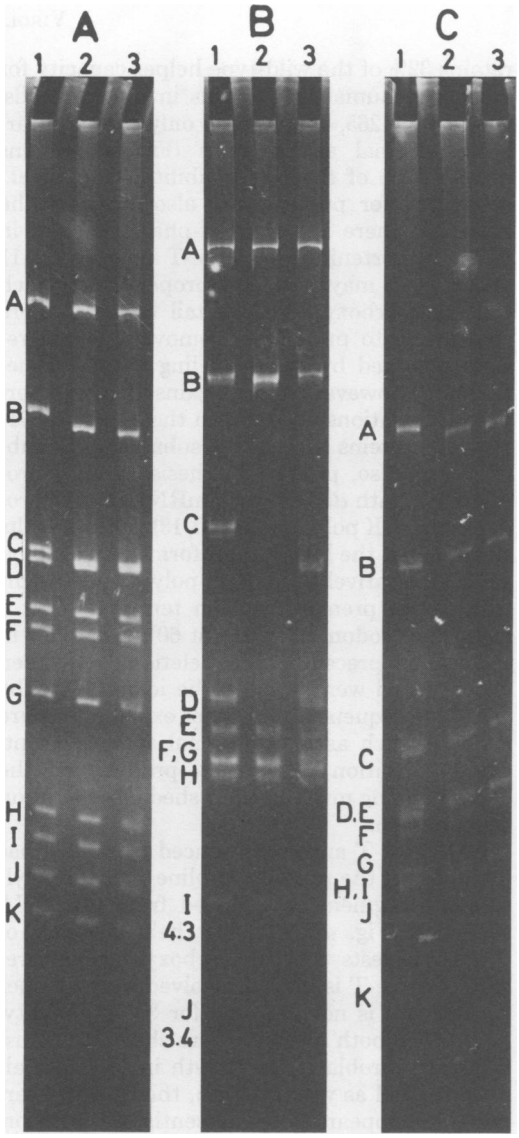

FIG. 1. *Mapping of the deletions in mutants dl1263 and dl1265. The purified DNA of each mutant was digested with restriction endonucleases HindII + III (A), HaeIII (B), or AluI (C), and electrophoresed on a 4% polyacrylamide slab gel overnight at 180 V. (1) Wild type; (2) dl1263; (3) dl1265.*

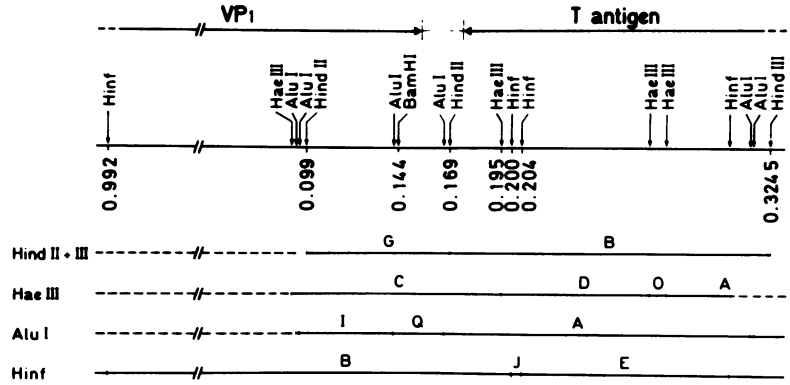

FIG. 2. *Location of some restriction enzyme cleavage sites in the SV40 DNA region encompassing the region coding for the carboxyl-terminal part of T antigen. Map units are based on the complete nucleotide sequence (6), with the unique EcoRI cleavage site as point zero.*

To map more precisely the deletion in *dl*1263 DNA, the *Hin*dII + III restriction fragments were 5′-terminally labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. *Alu*I digestion of the *Hin*d fragment B generated a singly 5′-terminally labeled subfragment, *Hin*d-B-A_1, which spanned the deletion (Fig. 2). This subfragment was partially digested with *Hin*f. Separation of the products showed the absence of the *Hin*f restriction site at map position 0.200 (data not shown). Hence, because *dl*1263 also lacks the *Hae*III site at position 0.195, the deletion must extend at least from position 0.195 to the *Hin*f site at position 0.200. To determine the exact size of the deleted DNA segment, the *Hin*f fragment between map positions 0.992 and 0.204 was 5′-terminally labeled and cleaved with *Hin*dII + III. Partial chemical degradation of the smaller subfragment by the procedure of Maxam and Gilbert (9) revealed an in-phase deletion of 33 base pairs (Fig. 3A and 4). This deletion corresponds to the absence of nucleotides E 2413 to E 2445 in the new numbering system adopted for the complete SV40 DNA sequence (6).

The exact extent of the deleted DNA segment in mutant *dl*1265 was determined as follows. *Hin*dII + III fragment B was 5′-terminally labeled and cleaved with *Hae*III; the nucleotide sequence of the resulting subfragment containing the deletion, *Hin*d-B-E_3, was then determined by partial chemical degradation. A deletion of 39 base pairs was found (Fig. 3B and 4), which corresponds to nucleotides E 2523 to E 2561. For *dl*1263, several gels were run independently, whereas for *dl*1265 concurrent information was obtained from both strands (not shown).

The 33-base-pair deletion in *dl*1263 does not account for the apparent 6,000-dalton reduction in size observed for the large-T antigen induced by this mutant (4). Considering this discrepancy, and also because the original mutant DNA stock showed some heterogeneity, the *dl*1263 mutant was recloned. The DNA obtained from the single-plaque-derived stock was perfectly homogeneous, and nucleotide sequence analysis completely confirmed the aforementioned in-phase deletion. This recloned *dl*1263 produced again a major 87K in addition to a minor 93K T antigen (Cole et al., submitted for publication). Therefore, it is unlikely that the discrepancy between the expected and the observed weight reduction for the large-T protein synthesized by *dl*1263 is due to the presence of a contaminant. The minor 93K polypeptide could correspond to the expected T antigen, which is shorter by 11 amino acids compared to the predicted sequence of wild-type large-T (Fig. 4). The fact that *dl*1263 retains 32% of the wild-type helper capacity for growth of human adenovirus in monkey cells, whereas *dl*1265, which lacks only the nine carboxyl-terminal amino acids (Fig. 4), retains merely 6.5% of the helper ability (Cole et al., submitted for publication), also supports the idea that there is no out-of-phase deletion in *dl*1263. Shortening the large-T antigen by 11 amino acids may affect the proper folding such that the carboxyl-terminal tail becomes more susceptible to proteolytic removal. Recent results obtained by pulse-labeling with [^{35}S]methionine, however, argue against a precursor-product relationship between the 93K and 87K large-T proteins (Cole et al., submitted for publication). Also, protein synthesis in vitro programmed with *dl*1263 early mRNA mainly produces the 87K polypeptide (10, 13). These results suggest that the latter is not formed by proteolysis. Alternatively, the 87K polypeptide might result from premature chain termination at a leaky stop codon, but at least 60 nucleotides in the region preceding the deletion have been checked and were found to be identical to the wild-type sequence. Still other explanations are possible, such as an additional splicing event. Characterization of the 87K protein and the corresponding mRNA might shed more light on this problem.

The large-T antigen produced by *dl*1265 has lost most of the peculiar, proline-rich carboxyl-terminal segment as deduced from the DNA sequence (Fig. 4). The low helper ability of *dl*1265 suggests that the carboxyl-terminal region of large-T is indeed involved in the helper process but is not required for SV40 viability. Also, since both mutants were shown to transform rat fibroblasts for growth in soft agar almost as well as wild type (2), the carboxyl-terminal tail appears to be inessential for transformation. It may be noteworthy that in *dl*1263 the nucleotides bordering on the deleted region are located within two loops of a possible secondary structure that can be constructed for the DNA segment encompassing the deletion (Fig. 5). Because *dl*1263 was prepared by S1 nuclease cleavage of a mixture of SV40 I and II DNAs (3), the aforementioned observation may be related to the hypothesis that supercoiling of a DNA molecule produces regions of interrupted secondary structure which migrate until they reach a segment where sufficient intrastrand complementarity is present (7, 21). No such model can be proposed for *dl*1265, because several possible secondary structures can be constructed for the relevant region.

We thank A. Van de Voorde and José Van der

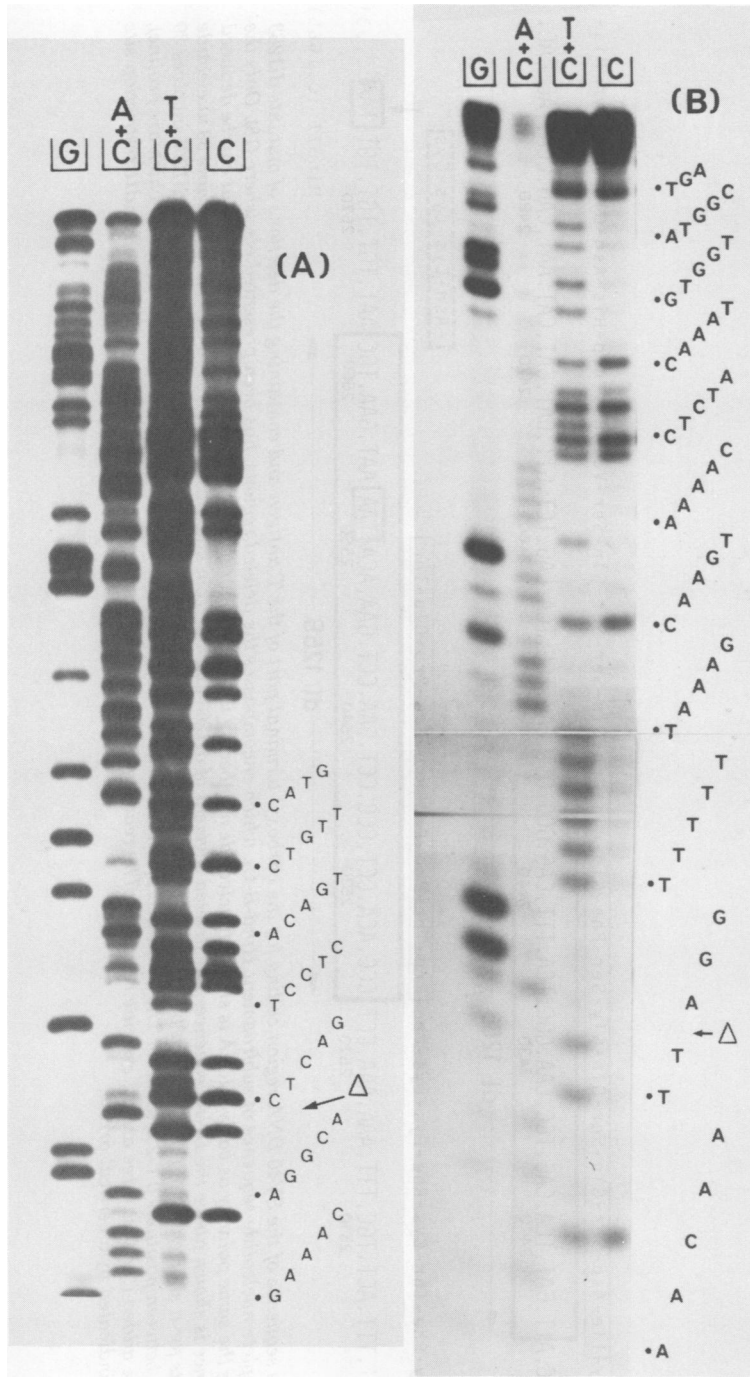

FIG. 3. *Sequence analysis of the relevant DNA segments of mutants dl1263 and dl1265. (A) The Hinf fragment between map positions 0.942 and 0.204 was 5'-terminally labeled and cleaved with HindII + III. The resulting singly 5'-terminally labeled subfragment containing the deletion was partially degraded according to the procedure of Maxam and Gilbert (9). The reaction products were fractionated on a 12% polyacrylamide slab gel containing 7 M urea and Tris-borate buffer (pH 8.3). The nucleotide specificity of each reaction is indicated above each lane. The site of the deletion is indicated by a delta symbol. (B) The subfragment containing the deletion in mutant dl1265 and labeled at only one 5' end was generated by kination of the HindII + III fragment B and subsequent cleavage with the restriction endonuclease HaeIII. Its sequence was determined as described in (A).*

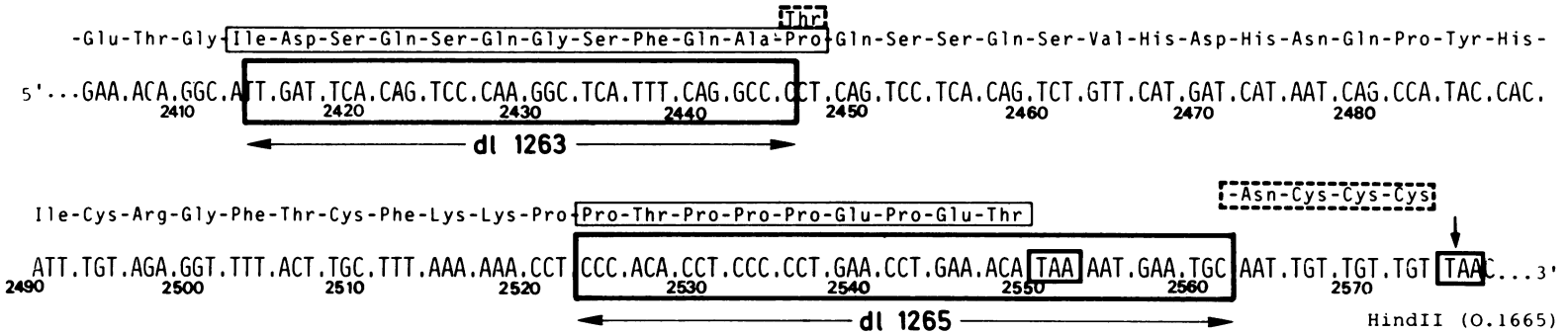

FIG. 4. *Nucleotide sequence of the SV40 DNA region coding for the carboxyl-terminal part of the T antigen and containing the deletions of mutants dl1263 and dl1265. The complete nucleotide sequence of subfragment Hind-B-C_2, which encompasses the deleted regions, has been presented elsewhere (18). Only the DNA strand that has the same polarity as early mRNA is shown (nucleotide numbering is taken from ref. 6), and the amino acid sequence that can be deduced from the DNA sequence is shown above the DNA sequence. The deleted segment of dl1263 extends from map positions 0.194 to 0.200 and contains 33 nucleotide pairs, which accounts for a shortening of the T antigen by 11 amino acids (12 amino acids removed; threonine is inserted). The deletion in dl1265 involves 39 base pairs and maps between positions 0.172 and 0.179. The T antigen of dl1265 lacks nine of the normal carboxyl-terminal residues (five of which are proline), but four new ones are added (the last three are all cysteine residues). The reading frame ends at another TAA termination codon. The HindII restriction site at position 0.1665 is indicated by a vertical arrow.*

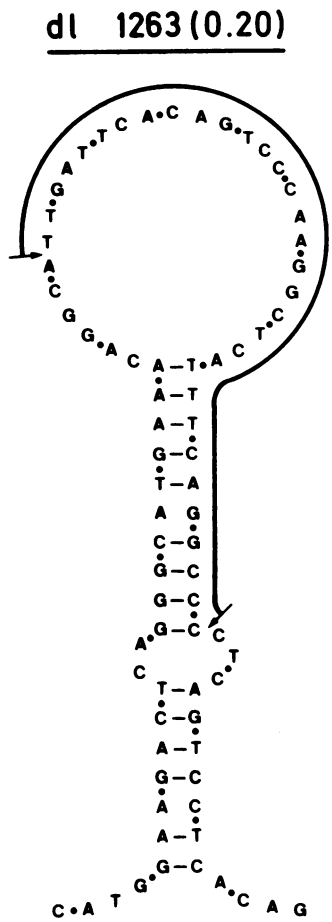

FIG. 5. *Possible secondary structure of the DNA segment encompassing the deletion in mutant dl1263. The deleted segment is 33 nucleotide pairs long and is marked with a heavy line. The bordering nucleotides (residues E 2412 and E 2446) are both located within a loop. The reading frame is marked off by dots, and base-pairing is indicated by dashes.*

Heyden for help with the biological aspects of this work, and L. Crawford for keeping us informed about his results. This research was supported by grants from the Algemene Spaar- en Lijfrentekas and from the Geconcerteerde Akties of the Belgian Ministry of Science.

LITERATURE CITED

1. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early messenger RNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:1274–1278.
2. **Bouck, N., N. Beales, T. Shenk, P. Berg, and G. diMayorca.** 1978. New region of the simian virus 40 genome required for efficient viral transformation. Proc. Natl. Acad. Sci. U.S.A. **75**:2473–2477.
3. **Carbon, J., T. E. Shenk, and P. Berg.** 1975. Biochemical procedure for production of small deletions in SV40 DNA. Proc. Natl. Acad. Sci. U.S.A. **75**:1392–1396.
4. **Cole, C., T. Landers, S. Goff, S. Manteuil-Brutlag, and P. Berg.** 1977. Physical and genetic characterization of deletion mutants of simian virus 40 constructed in vitro. J. Virol. **24**:277–294.
5. **Crawford, L. V., C. N. Cole, A. E. Smith, E. Paucha, P. Tegtmeyer, K. Rundell, and P. Berg.** 1978. Organization and expression of early genes of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:117–121.
6. **Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Voode, H. Van Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert.** 1978. Complete nucleotide sequence of SV40 DNA. Nature (London) **273**:113–120.
7. **Lebowitz, J., C. G. Garon, M. C. Y. Chen, and N. P. Salzman.** 1976. Chemical modification of simian virus 40 DNA by reaction with a water-soluble carbodiimide. J. Virol. **18**:205–210.
8. **Lebowitz, P., T. J. Kelly, D. Nathans, T. N. Lee, and A. M. Lewis.** 1974. A colinear map relating the simian virus 40 (SV40) DNA segments of six adenovirus-SV40 hybrids to the DNA fragments produced by restriction endonuclease cleavage of SV40 DNA. Proc. Natl. Acad. Sci. U.S.A. **71**:441–445.
9. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
10. **May, E., M. Kress, and P. May.** 1978. Characterization of two SV40 early mRNAs and evidence for a nuclear "prespliced" RNA species. Nucleic Acids Res. **5**:3083–3099.
11. **Paucha, E., R. Harvey, R. Smith, and A. E. Smith.** 1977. The cell-free synthesis of SV40 T antigens, p. 189–198. *In* P. May, R. Monier, and R. Weil (ed.), Early proteins of oncogenic DNA viruses. INSERM Colloquium, vol. 69. INSERM, Paris.
12. **Paucha, E., A. Mellor, R. Harvey, A. E. Smith, R. M. Hewick, and M. D. Waterfield.** 1978. Large and small tumor antigen from simian virus 40 have identical amino termini mapping at 0.65 map units. Proc. Natl. Acad. Sci. U.S.A. **75**:2165–2169.
13. **Paucha, E., and A. E. Smith.** 1978. The sequences between 0.59 and 0.54 map units on simian virus 40 DNA code for the unique region of small-t antigen. Cell **15**:1011–1020.
14. **Prives, C., E. Gilboa, M. Revel, and E. Winocour.** 1977. Cell-free translation of SV40 early messenger RNA coding for viral T-antigen. Proc. Natl. Acad. Sci. U.S.A. **74**:457–461.
15. **Tegtmeyer, P., M. Schwartz, J. K. Collins, and K. Rundell.** 1975. Regulation of tumor antigen synthesis by simian virus 40 gene *A*. J. Virol. **16**:168–178.
16. **Thimmapaya, B., B. S. Zain, R. Dhar, and S. M. Weissman.** 1978. Nucleotide sequence of the DNA template for the 3′ ends of SV40 mRNA. J. Biol. Chem. **253**:1613–1618.
17. **Tjian, R., G. Fey, and A. Graessmann.** 1978. Biological activity of purified simian virus 40 T-antigen proteins. Proc. Natl. Acad. Sci. U.S.A. **75**:1279–1283.
18. **Van Heuverswyn, H., A. Van de Voorde, and W. Fiers.** 1978. Nucleotide sequence of the SV40 DNA region coding for the carboxyl-terminal part of the T antigen. Eur. J. Biochem. **86**:335–344.
19. **Volckaert, G., A. Van de Voorde, and W. Fiers.** 1978. Nucleotide sequence of the simian virus 40 small-t gene. Proc. Natl. Acad. Sci. U.S.A. **75**:2160–2164.
20. **Walter, G., and H. Martin.** 1975. Simian virus 40-specific proteins in HeLa cells infected with nondefective adenovirus 2-simian virus 40 hybrid viruses. J. Virol. **16**: 1236–1247.
21. **Woodworth-Gutai, M., and J. Lebowitz.** 1976. Introduction of interrupted secondary structure in supercoiled DNA as a function of superhelix density: consideration of hairpin structures in supercoiled DNA. J. Virol. **18**:195–204.