JOURNAL OF VIROLOGY, Jan. 1979, p. 390-394
0022-538X/79/01-0390/05$02.00/0

Vol. 29, No. 1

Structural Relationship Between the 100,000- and 17,000-Molecular-Weight T Antigens of Simian Virus 40 (SV40) as Deduced by Comparison with the SV40-Specific Proteins Coded by the Nondefective Adenovirus Type 2-SV40 Hybrid Viruses

HAWLEY K. LINKE,† TONY HUNTER,* AND GERNOT WALTER

Tumor Virology Laboratory, The Salk Institute, San Diego, California 92112

Received for publication 11 July 1978

The two-dimensional peptide maps of the methionine-containing tryptic peptides of the 100,000-molecular-weight (100K) and 17K T antigens of simian virus 40 (SV40) have been compared. The two proteins share nine methionine-containing tryptic peptides in common. The 17K T antigen has two peptides not found in the 100K T antigen, and the 100K T antigen has 14 unique peptides. The peptide maps of the 100K and 17K T antigens were also compared with those of the SV40-specific proteins found in cells infected by the nondefective adenovirus type 2-SV40 hybrid viruses, which we have previously shown are encoded by defined sequences within the early region of SV40 (K. Mann, T. Hunter, G. Walter, and H. K. Linke, J. Virol. **24**:151-169, 1977). This comparison shows that the 100K and 17K T antigens share common *N*-terminal sequences coded for between 0.65 and 0.59 map units on the SV40 genome. Furthermore, none of the sequences in the 17K T antigen arises from the region between 0.54 and 0.18 map units. We deduce that the sequences unique to the 17K T antigen originate between 0.59 and 0.54 map units. This type of structural relationship between the 100K and 17K T antigens fits well with the proposed model (L. V. Crawford, C. N. Cole, A. E. Smith, E. Paucha, P. Tegtmeyer, K. Rundell, and P. Berg, Proc. Natl. Acad. Sci. U.S.A. **75**:117-121, 1978) for the expression of the early region of SV40.

Infection by simian virus 40 (SV40) induces the synthesis of at least two proteins early in infection before the initiation of viral DNA synthesis (1, 4, 5, 15, 17, 19, 22). These proteins have been defined by their ability to react with antisera directed against SV40-induced tumors and are known as T antigens. When separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, the larger T antigen migrates with an apparent molecular weight of 100,000 (100K T antigen), while the smaller protein migrates with an apparent molecular weight of 17,000 (17K T antigen). It has been shown that both these proteins are virally coded (10, 12, 15, 17), and that the 100K T antigen is required for the initiation of each round of viral DNA synthesis during productive infection (20). Both the 100K and 17K T antigens have been implicated in the initiation and/or maintenance of cellular transformation (3, 8, 19, 21).

Based on a number of different lines of biochemical (2, 6, 16) and genetic evidence (3, 19), Crawford et al. (5) proposed the following model for the expression of the early region of SV40. Both the 100K and 17K T antigens would share a common *N*-terminal sequence coded for between 0.65 and 0.59 map units. The coding sequence for the 17K T antigen would be continuous between 0.65 and 0.54 map units, where there are terminator codons in all three reading frames (6, 16). In the coding sequence for the 100K T antigen the sequences between 0.59 and 0.54 map units would be missing so that the *N*-terminal sequence from 0.65 to 0.59 map units would be joined to the region from 0.54 to 0.17 map units. This model has received support from the finding that the 100K and 17K T antigens do indeed share a common *N*-terminal sequence (13) and also have extensive amino acid sequences in common, while both possessing unique amino acid sequences (12, 14, 18). The finding that the larger early mRNA codes for the 17K T antigen while the smaller codes for the 100K T antigen is also consistent with this model (12).

† Present address: Molecular Biology Institute, University of California, Los Angeles, CA 90024.

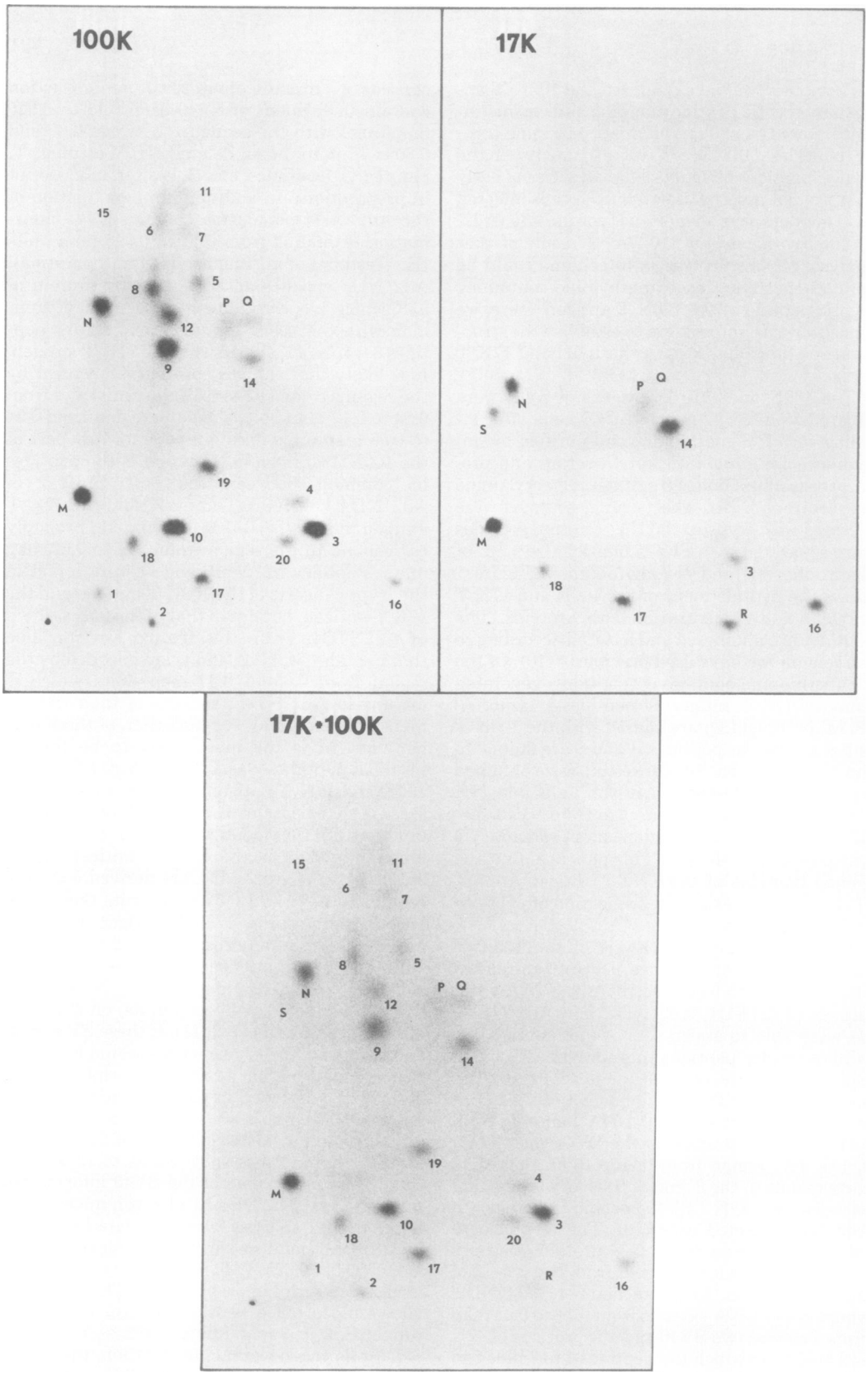
Previously we have analyzed the 100K T antigen by tryptic peptide mapping and can distinguish more than 20 methionine-containing tryptic peptides (10). We have also analyzed the tryptic peptides of an overlapping series of early SV40-coded proteins isolated from cells infected with nondefective adenovirus type 2-SV40 (Ad2-SV40) hybrid viruses (10). As a result of that analysis regions on the SV40 genome could be defined specifying each methionine-containing tryptic peptide of the 100K T antigen. Here, we have used this information to establish the structural relationship between the 100K and 17K T antigens.

The 100K and 17K T antigens of SV40 were isolated from SV40-infected TC7 cells after labeling with [^{35}S]methionine and purified by immunoprecipitation with anti-T serum and preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The purified proteins were digested with trypsin, and the resulting peptides were separated by a two-dimensional system of electrophoresis and chromatography. Figure 1 shows the peptide maps of the 100K and 17K T antigens and a mixture of both proteins. The peptides are numbered and lettered according to the scheme we have used previously (10). Of the 11 methionine-containing peptides generated from the 17K T antigen, 9 (peptides 3, 14, 16, 17, 18, M, N, P, and Q) are shared with the 100K T antigen, whereas peptides R and S are unique to the 17K T antigen. Similar results were obtained from the peptide analysis of the 100K and 17K T antigens synthesized in vitro from SV40-specific RNA in the mRNA-dependent reticulocyte lysate (data not shown). This indicates that the unique peptides of the 17K T antigen are unlikely to be the result of modification of peptides of the 100K T antigen.

From our previous analysis of the 100K T antigen of SV40 and a series of T antigen-related proteins encoded by a group of Ad2$^+$SV40 hybrid viruses (Ad2$^+$ND1, Ad2$^+$ND2, and Ad2$^+$ND4), we were able to assign sets of methionine-containing tryptic peptides to approximate coding regions within the early region of SV40 (10). For instance, Ad2$^+$ND1 is a nondefective Ad2 containing an insert of SV40 DNA mapping from 0.11 to 0.28 map units on the SV40 genome (7, 9, 11). The region from 0.28 to 0.17 map units corresponds to the 3′ end of the early region and encodes a 28K SV40-specific protein in Ad2$^+$ND1-infected cells (10). The tryptic peptides generated from the 28K protein (peptides 1 to 7) are therefore derived from the 3′ end of the early region. In a like manner, Ad2$^+$ND2, whose SV40 DNA extends from 0.11 to 0.44 map units, encodes two SV40-specific proteins of 42K and 56K from which the peptides 1 to 7 plus 8 to 11 are generated by tryptic digestion. Ad2$^+$ND4 consists of a mixture of viruses. A major fraction contains a segment of SV40 from 0.11 to 0.625 map units with the sequence between 0.60 and 0.50 map units being deleted (H. Westphal, T. Hunter, C. Lawrence, and G. Walter, manuscript in preparation). In addition a minor fraction of the Ad2$^+$ND4 population contains SV40 information extending from 0.11 to 0.625 map units (H. Westphal et al., manuscript in preparation). Ad2$^+$ND4 encodes an SV40-specific protein of 95K which has tryptic peptides 1 to 21 in common with the 100K T antigen and shares peptides 3, 14, 16, 17, and 18 with the 17K T antigen. It is likely that the 95K protein is encoded by the regions from 0.62 to 0.59 map units and from 0.54 to 0.17 map units since the region from 0.54 to 0.59 map units does not code for any part of the 100K T antigen (5). Therefore, the peptides in common between the 95K protein of Ad2$^+$ND4, the 100K T antigen, and the 17K T antigen, namely, 3, 14, 16, 17, and 18, probably correspond to the region from 0.59 to 0.62 map units. Peptides M, N, P, and Q were found in the tryptic digests of the 100K T antigen and the 17K T antigen, but not in that of the 95K protein of Ad2$^+$ND4. From this finding, we conclude that peptides M, N, P, and Q are encoded by the region from 0.65 to 0.62 map units which is missing in Ad2$^+$ND4. Analysis of their two-dimensional mobilities suggests that, of these four peptides, M is the most likely to be the *N*-terminal peptide, *N*-acetyl Met·Asp·Lys, of the 100K and 17K T antigens (13). Finally, one can assume that peptides R and S, which are found uniquely in the tryptic digest of the 17K T antigen and not in the 100K T antigen or the 95K protein of Ad2$^+$ND4, are derived from the region from 0.54 to 0.59 map units, known to encode part of the 17K T antigen (5). This information is summarized in Fig. 2.

In a previous publication, we assumed that Ad2$^+$ND4 only contained SV40 information from 0.11 to 0.59 SV40 map units, on the basis of published results (7, 9, 11). If these SV40 map end points had been correct, we would have had to conclude from our present data that the 17K T antigen was coded in part by a region downstream of 0.54 map units, which would have been in contradiction with the model of Crawford et al. (5). This inconsistency led us to reexamine the precise end points of the SV40 information integrated in Ad2$^+$ND4 by electron microscopic heteroduplex analysis. Our new data locate the most 5′ end point of early SV40 information in Ad2$^+$ND4 at 0.625 SV40 map units (H. Westphal et al., manuscript in preparation). This therefore allows us to place peptides 14, 16, 17, and 18 proximal to 0.59 map units.

Despite this new map information, the occurrence of peptide 3 in both the 100K and 17K T


100K
11
15
6
7
8
5
P
Q
N
12
9
14
19
4
M
10
3
18
20
1
17
16
2
17K
P
Q
N
S
14
M
3
18
17
R
16
17K•100K
15
11
6
7
8
5
N
P
Q
S
12
9
14
19
4
M
10
3
18
20
1
17
R
2
16

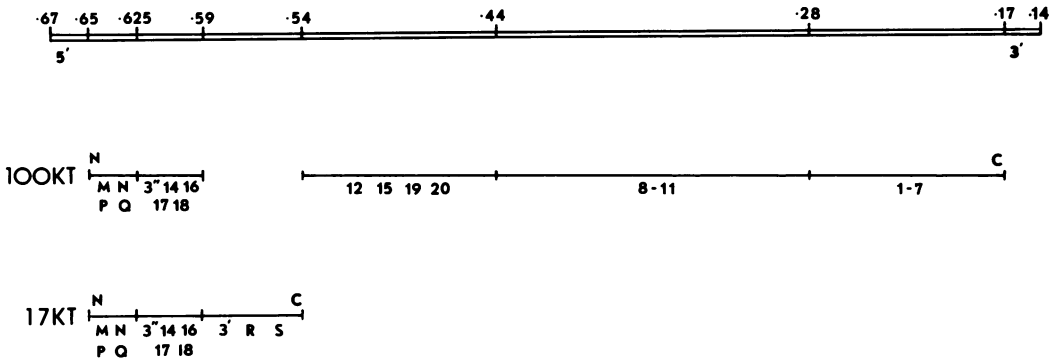


FIG. 2. *Assignment of methionine-containing tryptic peptides of the 100K and 17K T antigens to locations within the early region. The DNA of the early region of SV40 is represented by the open bar. The map coordinates of the 5' and 3' ends of the early cytoplasmic mRNAs (0.67 and 0.14), the 5' ends of the SV40 inserts of Ad2⁺ND1, Ad2⁺ND2, and Ad2⁺ND4 (0.28, 0.44, and 0.625), and the end points of the region in which viable deletions can occur (0.59 and 0.54) are shown. The locations of the methionine-containing tryptic peptides within the 100K and 17K T antigens are indicated as described in the text.*

antigens remained a puzzle, because we had previously located the coding region for peptide 3 near the 3' end of the early region (10). Perusal of the theoretical amino acid sequence of the 100K and 17K T antigens (6, 16) shows that the tryptic peptide Met·Lys is encoded by DNA sequences mapping at about 0.255 (peptide 3), 0.595 (peptide 3'), and 0.625 (peptide 3") map units. In addition, the migration of peptide 3 in our two-dimensional system is compatible with its being the dipeptide Met·Lys. Therefore the 100K T antigen would generate peptides 3 and 3", while the 17K T antigen would generate peptides 3' and 3". There is still an open question regarding peptide 16. Based on its migration in our system, we think it could be Met·Lys·Lys, a peptide coded for by sequences at 0.625 map units. This identity, however, would be inconsistent with the occurrence of this peptide in the 74K SV40-specific protein of Ad2⁺ND4, which we thought was coded downstream of 0.54 map units (10). This paradox may be resolved by our recent analysis of the structures of the SV40-specific RNAs present in the cytoplasm of Ad2⁺ND4-infected cells (H. Westphal et al., manuscript in preparation), which shows RNAs with unusual splices joining sequences at 0.60 map units to sequences at 0.44 map units within the SV40 early region. The 74K protein may be coded by one of these RNAs.

The work we have presented here provides independent support for the proposed model for the expression of the early region of SV40 (5). Our results confirm that the 100K and 17K T antigens share a common *N*-terminal region coded between the initiator AUG triplet at 0.65 map units (13) and 0.59 map units, whereas none of the sequences in the 17K T antigen arises from the region between 0.54 and 0.17 map units. Our data also show that the unique sequences of the 17K T antigen not in common with the 100K T antigen probably originate between 0.59 and 0.54 map units. One final conclusion can be drawn concerning the mRNA for the 95K protein in Ad2⁺ND4-infected cells. The almost complete identity of the peptides of the 95K protein with those of the 100K T antigen, coupled with the absence of peptides R and S in the 95K protein, strongly suggests that the mRNA for the 95K protein is lacking the sequences between

FIG. 1. *Tryptic peptide maps of [^{35}S]methionine-labeled 100K and 17K T antigens. Tryptic digests of [^{35}S]methionine-labeled 100K and 17K T antigens were prepared as previously described (10). In this case, one 14-cm dish of TC7 cells was infected with wild-type SV40 at a multiplicity of infection of about 20 PFU/cell. At 72 h postinfection at 37°C, the infected cells were labeled for 3 h with [^{35}S]methionine at 150 μCi/ml. The labeled cells were lysed at pH 8, and the cytoplasmic extract was immunoprecipitated with anti-T serum. The resultant immunoprecipitate was run on a preparative sodium dodecyl sulfate-polyacrylamide gel, which was subjected to autoradiography. The 100K and 17K T antigen bands were excised, and the extracted proteins were digested with trypsin. These digests were separated in two dimensions by electrophoresis at pH 4.7 toward the cathode with the origin on the left, followed by ascending chromatography. We used 6 × 10^3 cpm of the 100K T antigen digest and 2 × 10^3 cpm of the 17K T antigen digest. A mixture of the digests of the two T antigens was also run, containing half these amounts of radioactivity. The autoradiograms (Kodak NS5T X-ray film) were exposed for 30 days. The peptides have been lettered and numbered according to our previous scheme (10). In this instance, peptides 13 and 21 were not detectable, while peptide 15 is a very weak spot. These peptides have been found in variable yield on other occasions.*

0.59 and 0.54 map units and has undergone the same splicing event as the smaller of the two early SV40 mRNA's (2), which codes for the 100K T antigen (12).

This investigation was supported by Public Health Service grants no. CA-15365 and CA-17096 from the National Cancer Institute.

We thank the other members of the Tumor Virology Laboratory for their interest in this problem.

LITERATURE CITED

1. **Ahmad-Zadeh, C., B. Allet, J. Greenblatt, and R. Weil.** 1976. Two forms of simian virus 40-specific T-antigen in abortive and lytic infection. Proc. Natl. Acad. Sci. U.S.A. **73**:1097-1101.
2. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**: 1274-1278.
3. **Bouck, N., H. 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.
4. **Carroll, R. B., and A. E. Smith.** 1976. Monomer molecular weight of T antigen from simian virus 40-infected and transformed cells. Proc. Natl. Acad. Sci. U.S.A. **73**: 2254-2258.
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 SV40. Proc. Natl. Acad. Sci. U.S.A. **75**:117-121.
6. **Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Voorde, H. Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert.** 1978. The total nucleotide sequence of SV40 DNA. Nature (London) **273**:113-120.
7. **Kelly, T. J., Jr., and A. M. Lewis, Jr.** 1973. Use of nondefective adenovirus-simian virus 40 hybrids for mapping the simian virus 40 genome J. Virol. **12**: 643-652.
8. **Kimura, G., and R. Dulbecco.** 1973. A temperature-sensitive mutant of simian virus 40 affecting transforming ability. Virology **52**:529-534.
9. **Lebowitz, P. T., T. J. Kelly, Jr., D. Nathans, T. N. H. Lee, and A. M. Lewis, Jr.** 1974. A colinear map relating the simian virus 40 (SV40) DNA segments of six adenovirus-SV40 hybrids to the DNA fragments produced by restriction cleavage of SV40 DNA. Proc. Natl. Acad. Sci. U.S.A. **71**:441-445.
10. **Mann, K., T. Hunter, G. Walter, and H. K. Linke.** 1977. Evidence for simian virus 40 (SV40) coding of SV40 T-antigen and the SV40-specific proteins in HeLa cells infected with nondefective adenovirus type 2-SV40 hybrid viruses. J. Virol. **24**:151-169.
11. **Morrow, J. F., P. Berg, T. J. Kelly, Jr., and A. M. Lewis, Jr.** 1973. Mapping of simian virus 40 early functions on the viral genome. J. Virol. **12**:653-658.
12. **Paucha, E., R. Harvey, R. Smith, and A. E. Smith.** 1977. The cell free synthesis of SV40 T antigens. INSERM Colloq. **69**:189-198.
13. **Paucha, E., A. Mellor, R. Harvey, A. E. Smith, R. M. Hewick, and M. D. Waterfield.** 1978. Large and small tumor antigens from simian virus 40 have identical amino termini mapping at 0.65 map units. Proc. Natl. Acad. Sci. U.S.A. **75**:2165-2169.
14. **Prives, C., and Y. Beck.** 1977. Characterization of SV40 T-antigen polypeptides synthesized in vivo and in vitro. INSERM Colloq. **69**:175-188.
15. **Prives, C., E. Gilboa, M. Revel, and E. Winocour.** 1977. Cell free translation of simian virus 40 early messenger RNA coding for viral T-antigen. Proc. Natl. Acad. Sci. U.S.A. **74**:457-461.
16. **Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, C. L. Celma, and S. M. Weissman.** 1978. The genome of simian virus 40. Science **200**:636-646.
17. **Rundell, K., J. K. Collins, P. Tegtmeyer, H. L. Ozer, C.-J. Lai, and D. Nathans.** 1977. Identification of simian virus 40 protein A. J. Virol. **21**:636-646.
18. **Simmons, D. T., and M. A. Martin.** 1978. Common methionine-tryptic peptides near the amino-terminal end of primate papova virus tumor antigens. Proc. Natl. Acad. Sci. U.S.A. **75**:1131-1135.
19. **Sleigh, M. J., W. C. Topp, R. Hanich, and J. F. Sambrook.** 1978. Mutants of SV40 with an altered small t protein are reduced in their ability to transform cells. Cell **14**:79-88.
20. **Tegtmeyer, P.** 1972. Simian virus 40 deoxyribonucleic acid synthesis: the viral replicon. J. Virol. **10**:591-598.
21. **Tegtmeyer, P.** 1975. Function of simian virus 40 gene A in transforming infection. J. Virol. **15**:613-618.
22. **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.