JOURNAL OF VIROLOGY, Jan. 1983, p. 315–331
0022-538X/83/010315-17$02.00/0

Vol. 45, No. 1

# Improved Localization of Phosphorylation Sites in Simian Virus 40 Large T Antigen

FRANS VAN ROY, LUCIE FRANSEN, AND WALTER FIERS*
*Laboratory of Molecular Biology, State University of Ghent, B-9000 Ghent, Belgium*

Received 15 March 1982/Accepted 30 September 1982

The location of phosphorylation sites in the large T antigen of simian virus 40 has been studied both by partial chemical cleavage and by partial proteolysis of various forms of large T. These included the full-size wild-type molecule with an apparent molecular weight of 88,000, deleted molecules coded for by the mutants *dl*1265 and *dl*1263, and several shortened derivatives generated by the action of a cellular protease. These molecules differed from each other by variations in the carboxy-terminal end. In contrast, a ubiquitous but minor large T form with a molecular weight of 91,000 was found to be modified in the amino-terminal half of the molecule. In addition to the phosphorylation of threonine at position 701 (K.-H. Scheidtmann et al., J. Virol. **38**:59–69, 1981), two other discrete domains of phosphorylation were recognized, one at either side of the molecule. The amino-terminal region was located between positions 81 and 124 and contained both phosphothreonine and phosphoserine residues. The carboxy-terminal region was located between approximate positions 500 and 640 and contained at least one phosphoserine residue but no phosphothreonine. The presence in the phosphorylated domains of large T of known recognition sequences for different types of protein kinases is discussed, together with possible functions of large T associated with these domains.

Simian virus 40 (SV40) large tumor (T) antigen, or the viral A gene product, is a multifunctional protein with a central regulatory role in both productive infection and cell transformation (for a recent review, see reference 60). Protein phosphorylation is known to modulate the activity of multiple enzymes and regulatory proteins (27). Since large T antigen is a phosphoprotein (14, 48, 57, 63, 65), differential phosphorylation of its residues might yield molecular subclasses, each adapted to a particular set of functions and biological activities. Subclasses of large T with different isoelectric points have been described (42), and they may vary in phosphate content (21). Free monomers of large T antigen are reported to be underphosphorylated, whereas fast-sedimenting molecules in tetramer form or complexed with host-coded phosphoproteins of molecular weight ($M_r$) 48,000 to 55,000 (48K to 55K) are highly phosphorylated (15, 22, 34). Nevertheless, this conclusion has not been substantiated by the experiments of others (19). Also, the degree of phosphorylation is positively correlated with the nonspecific binding affinity of large T to calf thymus DNA (39) and with its affinity for host chromatin (22). However, it was also reported recently that in vitro dephosphorylation of T antigen at a number of specific phosphoserine (P-Ser) sites does not affect its binding to specific regions of SV40 DNA (50) and that dimer forms of large T antigen bind more specifically and tightly to SV40 DNA than do monomers or (highly phosphorylated) tetramers (3, 19). The most recent evidence comes from a study of proteins coded for by adenovirus 2-SV40 hybrid viruses and suggests that two regions of DNA-binding activity exist in large T: one contributes to nonspecific DNA binding, whereas the other is necessary for specific binding to the viral origin of replication (46). To obtain a better correlation between the phosphorylation state of large T and any of its functions, it seems imperative to locate and identify its phosphorylated residues.

We have previously described the presence of at least one P-Ser and one phosphothreonine (P-Thr) residue in an amino-terminal (N-terminal) 31K fragment generated from large T by mild digestion with *Staphylococcus* V8 protease (63). A truncated 33K large T molecule coded for by SV40 mutant *dl*1001 comprises this N-terminal 31K fragment. A complementary nonoverlapping carboxy-terminal (C-terminal) 58K fragment also contains at least one P-Ser and one P-Thr residue (63). The latter phosphorylation sites are present in the 42K and 56K large T variants coded for by the adenovirus-SV40 hybrid virus Ad2$^+$ND2, but large T of mutant

*dl*1265 lacks the C-terminal P-Thr site. In the present study, we have located these phosphorylation sites more precisely by both chemical degradation and partial proteolysis analysis.

## MATERIALS AND METHODS

**Cells and viruses.** CV-1 monkey cells were used throughout this study, except for a single experiment where SV40-transformed human cells (SV80) were extracted. These cells were grown in Dulbecco modified Eagle medium supplemented with 10% newborn calf serum and infected with SV40 as described (63). The wild-type (WT) SV40 used was strain 776, and mutant SV40 strains included the viable deletion mutants *dl*1263 and *dl*1265 (9; gift of C. N. Cole).

**Labeling and extraction of cells.** Cells were routinely labeled for 4 h at 42 h postinfection. Cells were labeled with [$^{35}$S]methionine (code SJ-204; Amersham International Ltd., England; 250 μCi in 1 ml/20 cm$^2$ of cells) after two washings in methionine-free medium supplemented with 2.5% dialyzed serum. For $^{32}$P labeling, cells were washed twice with phosphate-free medium supplemented with 2.5% dialyzed serum and then were incubated with 2 to 3 mCi of $^{32}P_i$ (code PBS-43; Amersham International) in 1 ml/20 cm$^2$ of cells. At the end of the labeling period, the cells were washed twice with ice-cold phosphate-buffered saline and then were extracted in situ for 0.5 h at 0°C with 400 μl of extraction buffer per 20 cm$^2$ of cells. This buffer consisted of 1% Nonidet P-40, 137 mM NaCl, 1 mM EDTA, 10% glycerol, 50 mM Tris-hydrochloride (pH 9.0), 500 Kallikrein inactivator units of aprotinin (Sigma Chemical Co., St. Louis, Mo.) per ml, 300 μg of phenylmethylsulfonyl fluoride (Serva, Heidelberg, Federal Republic of Germany) per ml, 150 μg of L-1-tosylamide-2-phenylethyl chloromethyl ketone (Serva) per ml, and 50 μg of *N*-α-*p*-tosyl-L-lysine chloromethyl ketone (Serva) per ml. In cases where proteolytic degradation products of large T were to be extracted, the latter protease inhibitors were omitted, and the pH value of the extraction buffer was decreased to 7.5. Lysates were scraped from the plates, cleared in an Eppendorf minicentrifuge, and either stored frozen or immunoprecipitated immediately.

**Large T purification.** Large T antigen was purified by immunoprecipitation with hamster anti-SV40 tumor serum generously provided by the Resources and Logistics Branch, National Cancer Institute, Bethesda, Md. (lot no. 79X000 128). Immune complexes were formed, adsorbed to protein A-Sepharose Cl-4B beads (Pharmacia Fine Chemicals, Uppsala, Sweden), washed, and eluted as described previously (63). In the case of large T antigen encoded by mutant *dl*1265, the eluted proteins were *S*-alkylated with *N*-ethylmaleimide (10). Routinely, large T proteins were analyzed on sodium dodecyl sulfate gels of 8% acrylamide and were calibrated with several marker proteins as described (63). If the proteins were to be processed further, the gels were dried immediately after electrophoresis with no washing or fixation step.

**Partial proteolysis peptide mapping.** Large T proteins were partially digested with *Staphylococcus aureus* V8 protease (Miles Laboratories, Slough, England) during reelectrophoresis on a 12.5% polyacrylamide gel by the method of Cleveland et al. (6). This one-dimensional analysis could be extended by loading a whole strip of this gel onto a new 12.5% gel and digesting it with increased concentrations of V8 during a second reelectrophoresis. Details of these one- and two-dimensional partial proteolysis mapping techniques, including the calibration peptides used, have been described (63). In this study all concentrations of V8 protease are given as the amount used per standard gel sample well (width, 6 mm).

**Peptide fingerprinting.** Proteins were eluted from preparative gel bands and were concentrated by precipitation with 20% trichloroacetic acid as previously described (63). Further treatment was done by the method of Smith et al. (53) and comprised oxidation with performic acid, digestion with trypsin, and analysis on thin-layer cellulose plates by electrophoresis at pH 1.9 in the first dimension and by chromatography in butanol-acetic acid-water-pyridine (30:6:24:20) in the second dimension.

**Peptide mapping by formic acid cleavage.** This chemical degradation was performed essentially by the method of Lam and Kasper (28). Protein bands (0.5 cm$^2$) were excised from dried gels and swollen by shaking for 1.5 h in 2 ml of 75% formic acid. Then the supernatant was removed by suction, and the wet gel fragments were incubated in closed vials overnight at 37°C without shaking. Thereafter, 3 ml of 1 M Tris base was added, and the gel pieces were shaken for 1 h. Finally, the gel pieces were incubated for 0.5 h in 0.5 ml of gel sample buffer. They were stored frozen until they were loaded in a sample well of a 12.5% polyacrylamide gel and electrophoresed.

**Phosphoamino acid analysis.** Proteins and proteolytic fragments thereof were eluted from dried gel pieces as described (63). Large T and its larger fragments (31K and above) were recovered by trichloroacetic acid precipitation as described (63). Smaller fragments (11K to 20K) were freed of sodium dodecyl sulfate and other salts efficiently and with high rates of recovery by gel filtration in a mixture of propionic acid, formic acid, and water (2:1:2, vol/vol/vol) by the method of Amons and Schrier (2). Acid hydrolysis was performed in 6 N HCl for 2 h at 110°C. Phosphoamino acids were separated by electrophoresis on cellulose thin-layer plates at pH 1.9 and were quantitated as described previously (63).

## RESULTS

**Different phosphopeptides in large T variant molecules.** We have previously found that proteolysis of undegraded WT large T antigen with small amounts of *S. aureus* V8 protease yields 31K and 58K phosphorylated intermediates originating from the N-terminal and C-terminal ends, respectively (63). These phosphopeptides are nonoverlapping, and both contain P-Ser as well as P-Thr residues. In this study we analyzed the generation of these diagnostic phosphopeptides in a number of species of large T, coded for by WT SV40 and by deletion mutants *dl*1265 and *dl*1263. WT SV40-infected cells yield undegraded 88K large T (referred to as form A), proteolytic degradation products with apparent $M_r$s of 83K and 78K (referred to as forms B and C), and minor amounts of a 91K species (referred to as

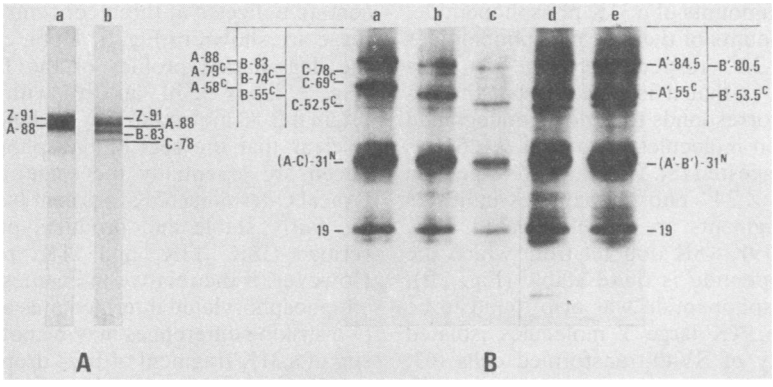


FIG. 1. Large T forms and their major *S. aureus* V8 proteolytic fragments. (A) $^{32}$P-labeled large T molecules extracted from SV40-infected cells either at pH 9.0 in the presence of protease inhibitors (lane a) or at pH 7.5 in the absence of inhibitors (lane b). Immunoprecipitates were analyzed on an 8% polyacrylamide gel. The different protein and peptide bands in this and the following figures are generally indicated by an identification code which is explained in the text and summarizes, if possible, the origin or nature of the protein, its apparent $M_r$ ($\times 10^3$) and, for peptide fragments, its location in the original undegraded molecule. If only a number is indicated, it refers to the apparent $M_r$ ($\times 10^3$) only. (B) Different $^{32}$P-labeled large T forms isolated as in (A) and then digested with V8 protease (10 ng per 6-mm sample well) during reelectrophoresis in a 12% acrylamide gel. The following large T polypeptides were analyzed: A-88 of WT SV40 (lane a), B-83 of WT SV40 (lane b), C-78 of WT SV40 (lane c), A'-84.5 of mutant *dl*1263 (lane d), and B'-80.5 of *dl*1263 (lane e). Outside rows of symbols correspond to fragments in outside lanes.

form Z) (Fig. 1A) (63). The genome of mutant *dl*1263 lacks 33 base pairs at map position 0.20 (61). Infection with this mutant yields apparently undegraded 84.5K large T (referred to as form A'), a proteolytic degradation product with an apparent $M_r$ of 80.5K (referred to as B'), and minor amounts of an 87.5K species (referred to as Z'). Large T from the mutant *dl*1265 lacks only the nine C-terminal amino acids, which are replaced by four new residues (61). Infection with this mutant yields a set of large T forms (referred to as A", B", C", and Z") that comigrate with the WT (A to Z) forms mentioned. In this study we will refer to any fragment of large T by indicating the large T species from which the fragment is derived, followed by the apparent $M_r$ (in thousands) and by the letters C, N, or I (superscript) to indicate whether the fragment resides N-terminally, C-terminally, or internally in the original molecule. For instance, according to this convention, mild V8 digestion of WT 88K large T yields the fragments A-31$^N$ and A-58$^C$ (see above).

We have now found that limited V8 proteolysis of most aforementioned $^{32}$P-labeled large T species produces WT-like N-terminal 31K fragments, but the C-terminal fragments differ from each other, as illustrated in Fig. 1B and summarized in Table 1. Thus, endogenous proteolysis during lytic growth or subsequent extraction clearly affects the C-terminal end of large T, and this is consistent with previous reports (33, 56). In contrast, digestion of both WT Z-91 and mutated Z'-87.5 molecules produced, in addi-

TABLE 1. Phosphoamino acid analysis of different large T antigen species and their major V8 proteolytic phosphopeptides

| Infecting virus | Large T species[a] | | | V8 phosphopeptides[b] | | |
|---|---|---|---|---|---|---|
| | Code[c] | Relative amt of[d]: | | Code[c] | Relative amt of[d]: | |
| | | P-Ser | P-Thr | | P-Ser | P-Thr |
| WT776 | Z-91 | 72 | 28 | Z-79$^C$ | ND[e] | ND |
| | | | | Z-58$^C$ | 67 | 33 |
| | | | | Z-34$^N$ | 73 | 27 |
| | A-88 | 74 | 26 | A-79$^C$ | ND | ND |
| | | | | A-58$^C$ | 70 | 30 |
| | | | | A-31$^N$ | 75 | 25 |
| | B-83 | 80 | 20 | B-74$^C$ | ND | ND |
| | | | | B-55$^C$ | 100 | 0 |
| | | | | B-31$^N$ | 72 | 28 |
| | C-78 | 82 | 18 | C-69$^C$ | ND | ND |
| | | | | C-52.5$^C$ | 100 | 0 |
| | | | | C-31$^N$ | 74 | 26 |
| *dl*1263 | Z'-87.5 | ND | ND | Z'-58$^C$ | ND | ND |
| | | | | Z'-34$^N$ | ND | ND |
| | A'-84.5 | 72 | 28 | A'-55$^C$ | 68 | 32 |
| | | | | A'-31$^N$ | 74 | 26 |
| | B'-80.5 | 79 | 21 | B'-53.5$^C$ | 100 | 0 |
| | | | | B'-31$^N$ | 73 | 27 |
| *dl*1265 | A"-88 | 82 | 18 | A"-58$^C$ | 100 | 0 |
| | | | | A"-31$^N$ | 73 | 27 |

[a] Data shown partially in Fig. 1A.
[b] Data shown partially in Fig. 1B and 2A.
[c] Explained in the text.
[d] Expressed as the percentage of the total counts for both phosphoamino acids.
[e] ND, Not done.

tion to minor amounts of a 31K phosphopeptide, significant amounts of the new phosphopeptides Z-34$^N$ and Z′-34$^N$, respectively (Fig. 2A). This finding indicates that it is unlikely that the A′-87.5 species corresponds to genuine undegraded large T antigen molecules of mutant *dl*1263, as previously suggested (8). Extended V8 digestion of A-31$^N$ and Z-34$^N$ phosphopeptides indicates that these fragments are closely related since both yield a 19K–18K doublet from which the 18K phosphopeptide is quite stable (Fig. 2B). The 34K phosphopeptide was also found to be typical of the 91K large T molecules isolated from a variety of SV40-transformed cells (63; data not shown).

Differences of 5K exist between large T forms A-88, B-83, and C-78, but apparently these differences drop to 2.5K to 4K if the N-terminal 31K fragment is cleaved off (yielding fragments A-58$^C$, B-55$^C$, and C-52.5$^C$) or if the deletion from the *dl*1263 mutant is introduced (yielding forms A′-84.5 and B′-80.5 and fragments A′-55$^C$ and B′-53.5$^C$). Large T, with a theoretical $M_r$ of 81,632 (60, 62), has an anomalously high apparent $M_r$ in sodium dodecyl sulfate-polyacrylamide gel (11, 23). Thus, it can be postulated that the discrepancies in $M_r$ mentioned above are due to a conformational anomaly in uncleaved and undeleted large T molecules. Previously, it has been found that deletions in the C-terminal region of large T do indeed cause an unexpectedly large drop in apparent $M_r$ (11, 17, 45).

We have previously reported that the large T antigen encoded by the mutant *dl*1265 lacks the C-terminal P-Thr site (63; Table 1). This is consistent with the definite localization of this phosphorylation site at position 701 by Scheidtmann et al. (48). As expected, this P-Thr residue is conserved in the A′-84.5 form of mutant *dl*1263 but is lost during endogenous proteolysis of the C-terminus (see the B-83, C-78, and B′-80.5 molecules in Table 1). However, this type of proteolysis apparently does not affect the phosphorylation of (all) C-terminal P-Ser residues, and neither does the *dl*1263 deletion.

**An N-terminal 10K fragment is not phosphorylated.** In addition to typical C-terminal 58K and N-terminal 31K and 34K fragments, a 79K phosphopeptide was generated when the A-88 and Z-91 species of WT large T were digested with small amounts (1 to 10 ng) of V8 protease (Fig. 1 and 2, Table 1). It is likely that this peptide arose by cleaving off a terminal 9K to 10K peptide from undegraded large T (see also Fig. 4A). Digestion of degraded large T species B-83 and C-78 yielded shortened equivalents of the 79K peptide (B-74 and C-69; Fig. 1, Table 1). This set of larger phosphopeptides was further analyzed by extended V8 proteolysis. Some representative digests at three concentrations of protease are shown in Fig. 3. Upon comparison of the degradation profiles of the B-74 peptides (Fig. 3, lanes b, b′, and b″) with those of the original B-83 molecules (Fig. 3, lanes e and e′), it is clear that the sets of phosphopeptides produced are essentially the same, including the typical C-terminal 55K fragment (see above) and the fairly stable end products previously described (18K, 13K, and 11K; reference 63). However, if the relative intensities of the different phosphorylated intermediates are compared, two striking differences may be noted: the intensity of a 31K fragment of B-74 drops, and that of a 19K fragment is considerably increased. Exactly the same relationship exists between the degradation patterns of C-69 peptides (Fig. 3, lanes c, c′, and c″) and C-78 molecules (Fig. 3, lanes d, d′, and d″) and between those of A-79 peptides and A-88 molecules (data not shown). Since the N-terminal 31K phosphopeptide is a precursor of a 19K–18K doublet (Fig. 2B; Fig. 3, lanes g, g′, and g″), we concluded that the A-79–B-74–C-69 set of peptides contains the C-terminus and arises by cleaving off by V8 protease from the respective precursor molecules an N-terminal 9K to 10K fragment. This is summarized schematically in Fig. 4A. Residual $^{32}P$ radioactivity in a 31K peptide of V8 digests of the molecules mentioned is most probably due to degradation of the fragments A-58$^C$, B-55$^C$,

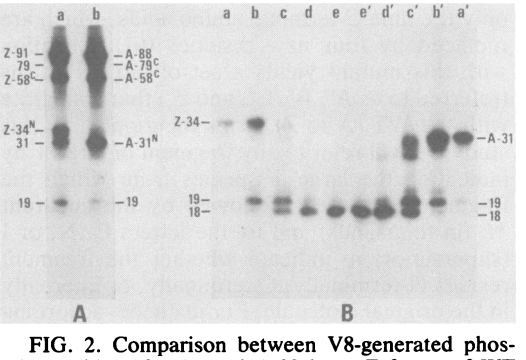


FIG. 2. Comparison between V8-generated phosphopeptides of Z-91 and A-88 large T forms of WT SV40. (A) $^{32}P$-labeled Z-91 (lane a) and A-88 (lane b) large T molecules, isolated as explained in the legend to Fig. 1 aged digested with 10 ng of V8 protease in an 8% polyacrylamide gel. (B) Phosphopeptides, excised from gel lanes similar to those depicted in panel (A) and subjected to extended proteolysis with increasing amounts of V8 protease during reelectrophoresis on a 12.5% polyacrylamide gel. Either Z-34 fragments (lanes a through e) or A-31 fragments (lanes a′ through e′) were analyzed. The following amounts of V8 were used: 0 ng (lanes a and a′), 10 ng (lanes b and b′), 100 ng (lanes c and c′), 1 μg (lanes d and d′), and 10 μg (lanes e and e′).

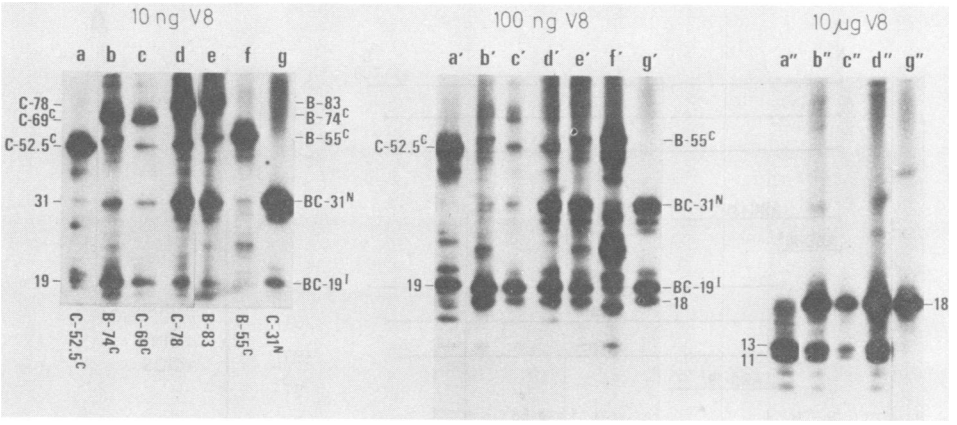


FIG. 3. Extended proteolysis of predominantly long V8-generated phosphopeptides of large T. Forms B-83 and C-78 of $^{32}$P-labeled WT-encoded T antigens were purified and digested with small amounts of V8 protease as described in the legend to Fig. 1. Selected peptide bands were excised and subjected to extended proteolysis with the amounts of V8 protease indicated at the top of the figure. The following fragments and proteins were analyzed as indicated at the bottom of the lanes: C-52.5$^{C}$ (lanes a, a', and a"); B-74$^{C}$ (lanes b, b', and b"); C-69$^{C}$ (lanes c, c', and c"); original C-78 molecules (lanes d, d', and d"); original B-83 molecules (lanes e and e'); B-55$^{C}$ (lanes f and f'); and C-31$^{N}$ (lanes g, g', and g"). All lanes came from a single electrophoretic run, but a shorter exposure time of the same gel was used for lanes e, e', and f.

and C-52.5$^{C}$, as illustrated by the control digests shown in Fig. 3 (lanes a, a', a", f, and f') and described in full detail further on.

The V8-generated 10K fragment presumed to be located N-terminally could not be detected by $^{32}$P labeling (Fig. 1B and 2A). However, when A-88 molecules labeled with [$^{35}$S]methionine were digested with smaller amounts of V8, a 10K fragment was easily detectable (Fig. 5A). Moreover, the same $^{35}$S-labeled unphosphorylated fragment was produced from the V8-generated A-31$^{N}$ peptide by extended V8 proteolysis but was absent from the comparable digests of the A-58$^{C}$ peptide (Fig. 5B). Thus, the suggestion that the A-31$^{N}$ peptide is processed to complementary 19K and 10K peptides seems well founded. Lanes c' and d' of Fig. 5B show also that most of the $^{35}$S radioactivity of A-31$^{N}$ is transferred to the 10K product. This finding also lends support to the N-terminal position of the latter because the N-terminal region of large T is known to be extremely rich in methionine residues (60, 62). On the other hand, all of the $^{32}$P radioactivity in A-31$^{N}$ is transferred to the 19K–18K doublet (Fig. 2B and Fig. 3, lane g").

The origin of the V8-generated 10K peptide was further explored by preparative digestion of $^{35}$S-labeled A-88 molecules, prepared this time from SV80 cells. The 10K peptide, together with undegraded A-88 large T, was eluted and analyzed by trypsin digestion and two-dimensional peptide mapping (Fig. 6). The lower half of the peptide map of A-88 molecules was particularly closely comparable with maps of large T tryptic peptides reported previously (43, 53, 54). The upward pointing arrows in Fig. 6 indicate peptides also present in the digest of the 10K fragment. This was especially evident when both maps were calibrated by the ninhydrin-stained peptide profile of bovine serum albumin, added to the digests as a carrier protein. Moreover, the peptides of the 10K fragment corresponded to the set of [$^{35}$S]methionine tryptic peptides shared by large T and small t (53, 54) and, most interestingly, included the spot known to be the N-terminal tryptic peptide N-acetyl-Met-Asp-Lys (labeled with an N in Fig. 6) (38, 54). These data provide definitive proof that the unphosphorylated 10K peptide (A-10$^{N}$) contains the N-terminus. As expected, all peptides of A-10$^{N}$ are also present in a tryptic digest of V8-generated A-31$^{N}$, whereas the tryptic digest of A-58$^{N}$ lacks the N-terminal peptide mentioned but yields a peptide previously reported to be present in the C-terminal half of large T (indicated by the downward arrow in Fig. 6; further data not shown) (43, 54).

**An internal 12K fragment at position 135 to 239 is not phosphorylated.** According to the nucleic acid sequence of the SV40 A gene (60, 62), large T contains two aspartyl-prolyl bonds, at positions Asp$^{134}$-Pro and Asp$^{239}$-Pro. Such bonds are known to be particularly acid labile (29). Thus, we have subjected the A-88, B-83, and C-78 forms of $^{32}$P-labeled large T antigen to partial formic acid cleavage, the better to localize the phosphorylation sites. Although nonspecific degradation also occurred to some extent, a

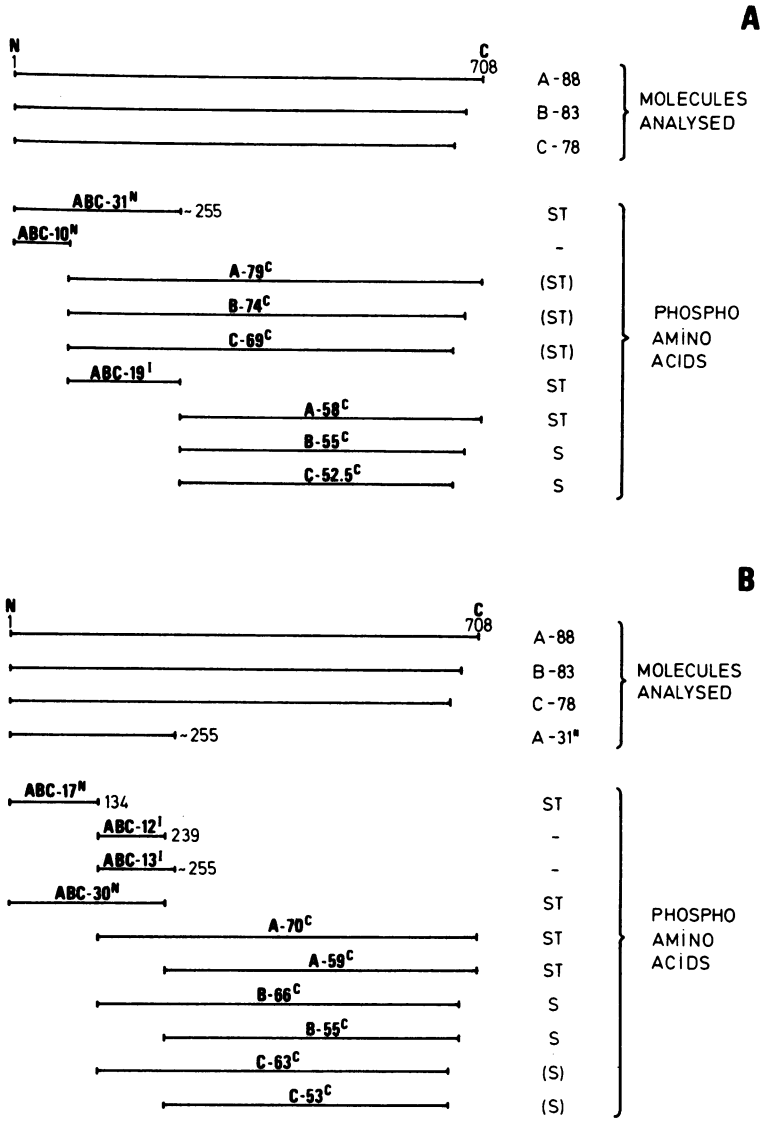


FIG. 4. Schematic representation of the cleavage sites in SV40 large T antigen. N and C mark the N- and C-terminals, respectively. Numbers refer to positions in the amino acid sequence (60, 62). Peptides are designated by a code explained in the text. The phosphoamino acid composition of the different fragments is shown at the right: S and T stand for P-Ser and P-Thr, respectively. Letters in parentheses indicate that the given phosphoamino acid composition is likely but has not been confirmed experimentally. (A) Peptides yielded by V8 protease digestion; see Table 1 and Fig. 3, 5, and 6 for most of the data. (B) Peptides yielded by limited acid hydrolysis; see Table 2 and Fig. 7.

number of prominent phosphorylated fragments could be seen (Fig. 7, lanes a through d). These could be identified as indicated in Table 2 and the scheme of Fig. 4B. It follows from the data in Table 2 that all phosphopeptides obtained from A-88 molecules contain both P-Ser and P-Thr residues, but the 66K and 55K peptides of the B-83 molecule lack P-Thr. This confirms the C-terminal localization of the latter peptides, which lack the P-Thr$^{701}$ site. In addition, the A-31$^{N}$ fragment of V8-cleaved large T yielded a single 17K phosphorylated fragment upon acid treatment (Fig. 7, lane e). This confirms the localization of the 17K peptide in the N-terminal half of large T (Fig. 4B). No 12K phosphopeptide was detectable in any of the digests shown

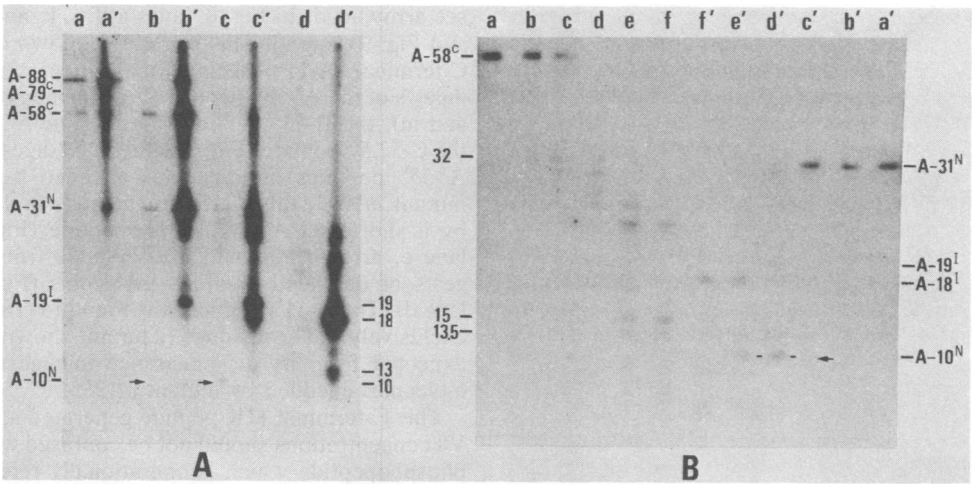


FIG. 5. Partial proteolysis peptide mapping of [$^{35}$S]methionine-labeled large T molecules. Large T molecules purified by immunoprecipitation and electrophoresis on an 8% gel as described in the legend to Fig. 1. The A-88 protein bands were excised and subjected to partial proteolysis with V8 protease during reelectrophoresis on a 12.5% polyacrylamide gel (A). Either $^{35}$S-labeled large T (lanes a, b, c, and d) or $^{32}$P-labeled large T (lanes a', b', c', and d') were digested with the following amounts of V8: 1 ng (lanes a and a'), 10 ng (lanes b and b'), 100 ng (lanes c and c'), and 1 μg (lanes d and d'). (B) $^{35}$S-labeled peptides of the A-88 large T form, excised from gel lanes similar to lane b of (A) and subjected to extended proteolysis with increasing amounts of V8 protease during reelectrophoresis on a 12.5% polyacrylamide gel. Either A-58$^{C}$ (lanes a through f) or A-31$^{N}$ (lanes a' through f') peptides were analyzed. The following amounts of V8 were used: 0 ng (lanes a and a'), 1 ng (lanes b and b'), 10 ng (lanes c and c'), 100 ng (lanes d and d'), 1 μg (lanes e and e'), and 10 μg (lanes f and f'). The arrows and bars point out a nonphosphorylated 10K peptide (A-10$^{N}$) discussed in the text.

TABLE 2. Phosphoamino acid analysis of large T antigen peptides generated by cleavage at Asp-Pro bonds

| Large T form | Peptide location[a] | $M_r$ (K) | | Relative amt of[b]: | |
|---|---|---|---|---|---|
| | | Theoretical[c] | Observed[d] | P-Ser | P-Thr |
| A-88 | N-I-C | 81.6 | 88 | ++ | + |
| | I-C | 66 | 70 | ++ | + |
| | C | 54 | 59 | ++ | + |
| | N-I | 27.5 | 30 | ++ | + |
| | N | 15.5 | 17 | ++ | + |
| | I | 12 | U[e] | − | − |
| B-83 | N-I-C | ~80 | 83 | +++ | + |
| | I-C | ~64 | 66 | + | − |
| | C | ~52 | 55 | + | − |
| | N-I | 27.5 | 30 | ++ | + |
| | N | 15.5 | 17 | ++ | + |
| | I | 12 | U | − | − |

[a] N, N-terminal; I, intermediately positioned; C, C-terminal.

[b] The relative abundance of each phosphoamino acid was evaluated from relative spot intensities on autoradiographs and is represented by the number of + signs. −, Undetectable spots.

[c] Calculated from the amino acid sequence of large T (60, 62), assuming cleavage at the Asp$^{131}$-Pro and the Asp$^{239}$-Pro bonds.

[d] Data shown in Fig. 7, lanes a through c.

[e] U, Undetectable by $^{32}$P labeling.

in Fig. 7. Therefore, the 12K fragment in the intermediate position (from Pro$^{135}$ to Asp$^{239}$) is apparently not phosphorylated.

**Improved localization of C-terminal phosphorylation sites.** C-terminal fragments with $M_r$s between 58K and 52.5K were isolated by limited V8 proteolysis from several T antigen forms (Fig. 1B, Table 1). They were further analyzed by extended V8 proteolysis (Fig. 8 and 9). Phosphopeptides typical of each large T species or subsets of these species (indicated by dots in the figures) as well as phosphopeptides common to all species were revealed. Most of the precursor-product relationships within these groups of phosphorylated fragments could be worked out from or confirmed by an extensive setup of two-dimensional V8 proteolytic degradations. The details of this procedure have already been published (63), and a few typical examples at high and moderate V8 concentrations in the second dimension are given here (Fig. 10). Thus, by knowing the structure peculiar to each form of large T (see above) we were able to construct a best-fitting model for the location of these phosphopeptides in the large T molecule (Fig. 11). For the interpretation of the degradation patterns, it should be borne in mind that only phosphorylated fragments are detected. Prominent phosphopeptides in the cleavage patterns

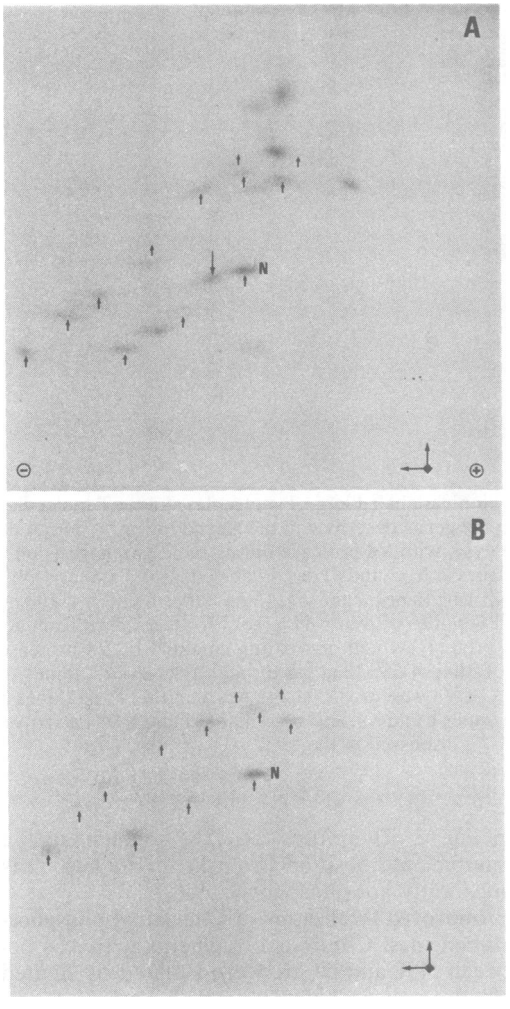


FIG. 6. Tryptic peptide fingerprints of the [$^{35}$S]methionine-labeled A-88 form of large T (A) and of A-10$^{N}$, a 10K fragment thereof (B), prepared as described in the legend to Fig. 5. These molecules were digested with trypsin, and the resulting peptides were analyzed on cellulose thin-layer chromatography plates by electrophoresis at pH 1.9 towards the cathode (on the left) followed by ascending chromatography in the second dimension. Upward-pointing arrows indicate spots shared by A-88 and A-10$^{N}$ molecules. N indicates the N-terminal peptide N-Acetyl-Met-Asp-Lys (38, 54). The downward arrow indicates a peptide present in the C-terminal half of large T (see text).

(Fig. 11, horizontal arrowheads) either bear major phosphorylation sites or arise because some peptide bonds are cleaved preferentially (Fig. 11, black vertical arrows).

Most worthy of interest are the smaller phosphopeptides and their mutual relationships. A stable 11K phosphopeptide is cleaved off at rather low enzyme concentrations (10 to 100 ng; see arrowheads in Fig. 8, lanes a, f, i, k, and m, and Fig. 9, lane a). This peptide is known to be C-terminal (A-11$^{C}$) because it disappears from digests of the A″-58$^{C}$ peptide (Fig. 8, lanes g, j, l, and m), the B-55$^{C}$ peptide (Fig. 8, lane b), and the C-52.5$^{C}$ peptide (Fig. 9, lane b). In digests of A′-55$^{C}$ peptides of molecules encoded by the mutant *dl*1263, this A-11$^{C}$ fragment is replaced by a shortened A′-7$^{C}$ phosphopeptide (Fig. 9, lane c, arrowhead), which disappears from digests of degraded B′-53.5$^{C}$ molecules (Fig. 9, lane d). The A-11$^{C}$ peptide was found to contain exclusively P-Thr residues (data not shown), as expected from its disappearance in digests of molecules encoded by mutant *dl*1265.

This C-terminal 11K peptide generated at low V8 concentrations should not be confused with a phosphopeptide which is operationally referred to as a 12K peptide but actually almost comigrating with the A-11$^{C}$ fragment. This comigration complicates the interpretation of the two-dimensional proteolytic patterns. The 12K peptide with internal location was found in the digests of all types of molecules examined, but only at the higher V8 concentrations (1 to 10 μg; Fig. 8, lanes d, l, and n, and Fig. 9, lanes f, g, and h). The latter peptide is related to an internal 13K phosphopeptide which is also found in digests of all types of large T molecules. Indeed, extended cleavage of a 17K V8 phosphopeptide common to all forms of large T yields a 16K phosphopeptide as well as those of 15K, 13K, and 12K, whereas cleavage of this 16K peptide yields phosphorylated 14K and 12K but not 15K or 13K products (Fig. 10A). Similar arguments can be proposed from the extended V8 degradation of 19K–18K phosphopeptides and other doublet fragments (see also Fig. 11). Both internal 12K and 13K phosphopeptides (Fig. 8 and 9, horizontal bars) can be processed further (at least partially) to a 7K and a 4K phosphopeptide at a very high V8 concentration (40 μg; data not shown), but fine mapping of the latter very small peptides has not so far been possible. The 13K phosphopeptide was found to carry exclusively one or more P-Ser residues (data not shown).

Since the *dl*1263 deletion which removes residues 663 to 674 from large T (61), does not affect the internal 12K–13K doublet fragments (Fig. 9, lane g), the latter peptides must be separated from the C-terminus by a distance of at least 5K. Alternatively, the major internal 13K peptide is related to major 17K, 19K, and 43K phosphopeptides (Fig. 10 and 11). The latter peptides are common to all large T forms examined and can be derived, for example, from major A-27, A-29, and A-53 peptides merely by cleaving off the A-11$^{C}$ peptide each time. These arguments lend strong support to the idea that the A-11$^{C}$ and A-13$^{I}$ phosphopeptides are neighboring nonover-

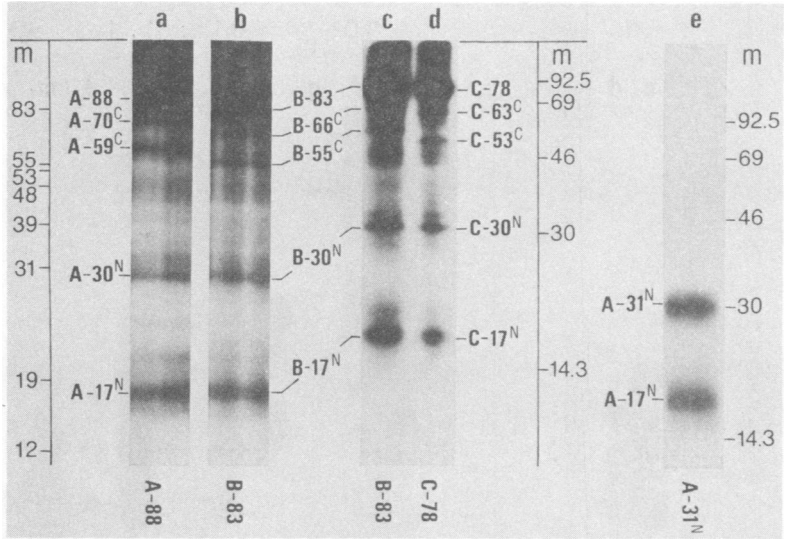


FIG. 7. Formic acid cleavage of large T molecules and a V8-generated fragment thereof. $^{32}$P-labeled A-88, B-83, and C-78 large T forms and the A-31$^N$ phosphopeptide were purified as described in the legend to Fig. 1. These molecules were cleaved by formic acid hydrolysis as described in the text, and the reaction mixtures were analyzed on a 12.5% polyacrylamide gel. In a first electrophoretic run, phosphopeptides of A-88 (lane a) and B-83 (lane b) large T molecules were analyzed as indicated at the bottom of the lanes. In a second run, phosphopeptides of B-83 (lane c) and C-78 (lane d) molecules were analyzed. Phosphopeptides of the A-31$^N$ fragment (lane e) were separated in a third run. The letter m indicates for each run a lane of $M_r$-markers. In the first run, these consisted of the following molecules: undegraded B-83 and V8-generated peptides of this large T form (55K, 31K, and 19K), human nonviral T or p53 (53K), and V8-generated peptides of this cellular protein (48K, 39K, and 12K; see reference 63). In the other two runs, commercial calibration peptides were used (63).

lapping fragments. Furthermore, from the analysis of the shorter phosphopeptides unique to each of the different forms of large T, it can be deduced that an A-23$^C$ peptide is the shortest one comprising both the nonoverlapping A-11$^C$ and A-13$^I$ phosphopeptides. It should also be noted here that under certain experimental conditions, C-terminal phosphopeptides appear to display different conformational states. Thus, the A-23$^C$, A-11$^C$, and A′-7$^C$ fragments can also be present as 21K, 9K, and 5.5K conformers, respectively (indicated by brackets in Fig. 8, lanes a and c, and in Fig. 9, lanes a, c, e, and g). The latter peptides are not proteolytic products of the former since they are not present consistently in extended digestions (Fig. 8, lanes i and k) and since they comigrate with the slower migrating forms upon reelectrophoresis (see also Fig. 10).

We have shown before that the *dl*1265 deletion removes all P-Thr residues from the A″-58$^C$ fragment (63). Here it was found that the A-11$^C$ fragment contains this P-Thr residue(s) but lacks any additional P-Ser residues. Previously (63) and here (Fig. 3 and 8), we have also shown that upon cleavage of A-58$^C$ and equivalent C-terminal fragments of other large T forms with high concentrations of V8, almost all incorporated $^{32}$P radioactivity is transferred to quite stable phosphopeptides of 11K to 12K and of 13K. Therefore, it is very likely that the 13K–12K doublet contains at least the predominant and possibly all P-Ser residues present in the C-terminal region of large T. Making allowance for conformation-specific influences on apparent $M_r$, we can conclude so far that the phosphorylation of serine residues in the C-terminal region of large T is restricted to a region between approximate positions 500 and 640 (Fig. 11 and 12).

## DISCUSSION

The present state of our knowledge about the phosphorylation sites in large T antigen is represented in Fig. 12. Taking the N-terminus as the starting point, we have detected a 10K peptide which contains this terminus but lacks phosphorylation sites. The length of this peptide corresponds roughly to that of the N-terminal segment of 82 amino acid residues common to the large T and small t antigens (60). The small t antigen has never been found to be phosphory-

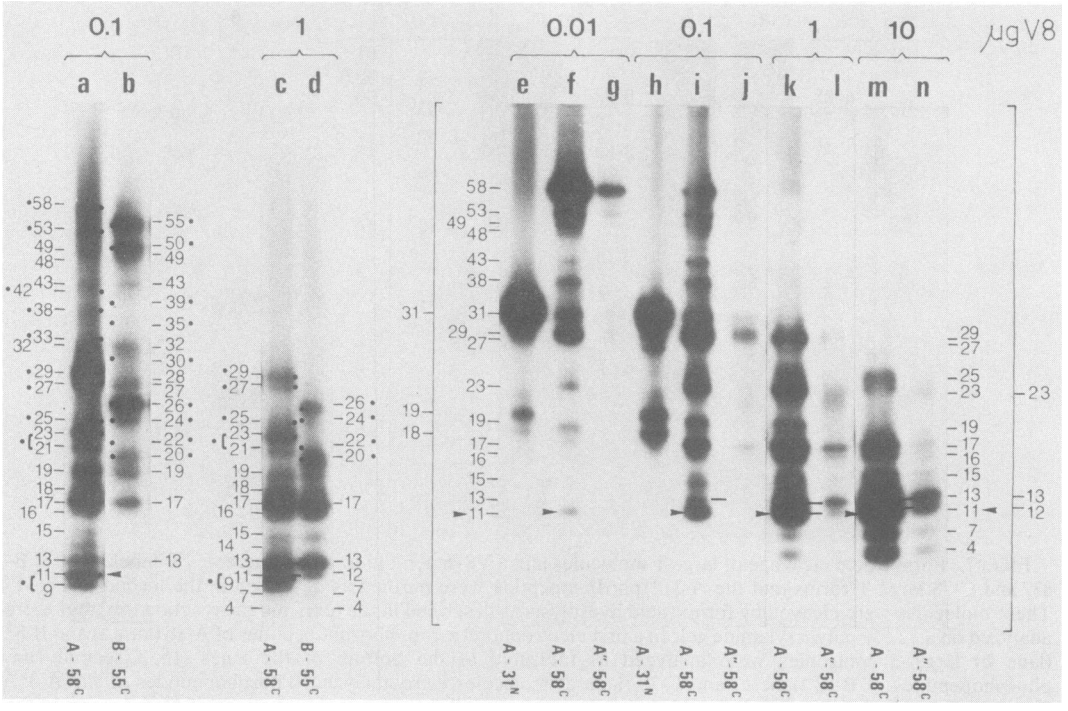


FIG. 8. Extended proteolysis of predominantly C-terminal V8-generated phosphopeptides of large T. Different forms of $^{32}$P-labeled large T antigens were purified and digested with low amounts of V8 protease as described in the legend to Fig. 1B. Selected peptide bands were excised and subjected to extended proteolysis with the amounts of V8 protease indicated at the top of the figure. The following fragments (indicated at the bottom of the lanes) were analyzed: A-58$^C$ of WT-encoded 88K large T (lanes a, c, f, i, k, and m); B-55$^C$ of WT-encoded 83K large T (lanes b and d); A-31$^N$ of WT-encoded 88K large T (lanes e and h); and A"-58$^C$ of 88K large T encoded by mutant *dl*1265 (lanes g, j, l, and n). Outside rows of numbers correspond to apparent $M_r$s of fragments in outside lanes. All other symbols are explained in the text. Lanes a through d and lanes e through n came from two different gel runs.

lated in vivo (32, 49, 55, 67). This does not necessarily mean that the same fragment could not be phosphorylated in large T, because the N-terminal ends of small t and large T are known to have different conformations (66). Partial acid hydrolysis of large T produced an N-terminal 17K fragment that extended from position Met$^1$ to Asp$^{134}$ and contained both P-Ser and P-Thr residues. This finding agrees with the characterization by Schwyzer et al. (49) of a 17K phosphorylated N-terminal tryptic peptide of large T, which would extend up to position Arg$^{130}$. An internal 12K fragment which extends from position Pro$^{135}$ to Asp$^{239}$ was apparently not phosphorylated.

Taken together, these data lead to the conclusion that phosphorylation in the N-terminal region of large T of both Ser and Thr residues ought to be located between positions Thr$^{81}$ and Thr$^{124}$. Indeed, according to the nucleic acid sequence of the SV40 A gene (60, 62), Thr$^{81}$ is the first Ser or Thr residue after Thr$^{57}$ which definitely resides in the unphosphorylated N-terminal 10K fragment characterized. On the other hand, Thr$^{124}$ is the last Ser or Thr residue before Pro$^{135}$ which is the N-terminal end of the internal 12K fragment found to be unphosphorylated. Previously we have shown that V8-cleavage of a 33K large T fragment, encoded by the mutant *dl*1001 and extending from the N-terminus to position 272, yields the typical N-terminal 31K phosphopeptide (63). This implies that the A-31$^N$ fragment extends to about position 255. This leaves a short stretch of about 15 residues, starting at position Pro$^{240}$, about which we have no direct information. This stretch contains phosphorylatable Ser residues at positions 242 and 247, and it contains no Thr. However, it is very unlikely that these residues represent a major site(s) of phosphorylation since formic acid hydrolysis of the A-31$^N$ fragment yields only the A-17$^N$ phosphopeptide and no detectable $^{32}$P-labeled 12K to 13K peptide with a presumed location between approximate positions 135 and 255 (see also Fig. 4B and Fig. 7, lane e). In addition, Schwyzer et al. (49) charac-

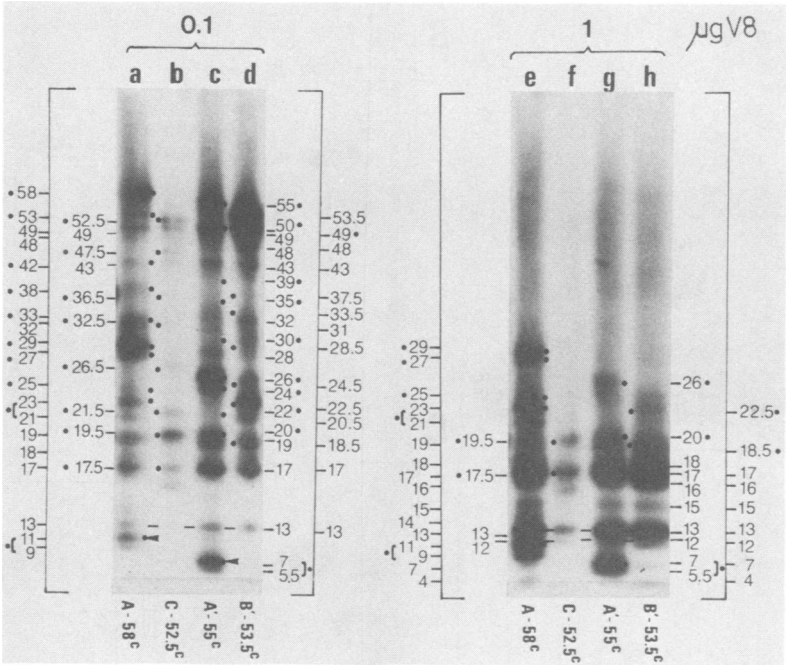


FIG. 9. Extended proteolysis of C-terminal V8-generated phosphopeptides of large T. Selected $^{32}$P-labeled peptides were subjected to extended V8 proteolysis as described in the legend to Fig. 5. The following fragments were analyzed: A-58$^C$ of WT-encoded 88K large T (lanes a and e); C-52.5$^C$ of WT-encoded 78K large T (lanes b and f); A'-55$^C$ of 84.5K large T encoded by mutant *dl*1263 (lanes c and g); and B'-53.5$^C$ of *dl*1263-encoded 80.5K large T (lanes d and h). Outside rows of numbers correspond to apparent $M_r$s of fragments in outside lanes. The other symbols are explained in the text. All lanes came from the same gel run.

terized a phosphate-free internal 40K tryptic peptide with N-terminus at position Lys$^{131}$, which should exclude phosphorylation of the residues between position 240 and 255 (see also Fig. 12).

Our localization of the C-terminal phosphorylation sites in large T was largely based on partial proteolysis maps of different large T forms. These forms were shortened internally near the C-terminus by the *dl*1263 deletion (form A'), at the C-terminus itself by the *dl*1265 deletion (form A''), or by the action of a cellular protease (forms B, B', B'', C, and C''). This procedure enabled us to locate a 13K phosphopeptide internally between approximate positions 500 and 640 and carrying one or more P-Ser residues. This confirms and extends the results of Schwyzer et al., mentioned above (49). Since the exceptional conformation of the C-terminus of large T makes accurate determination of the $M_r$s of C-terminal peptides difficult, we have so far been unable to improve our localization of the P-Ser site(s) in that region, which contains a total of 10 phosphorylatable Ser residues. This 13K phosphopeptide is separated from the C-terminus by a P-Thr-containing fragment with an apparent $M_r$ of 11K. The *dl*1265 mutation removes all radioactivity from this A-11$^C$ peptide (this paper) as well as all P-Thr residues from the A-58$^C$ fragment (63), so that either Thr$^{701}$ or Thr$^{708}$ carries this phosphate. Recently, Scheidtmann et al. (48) have mapped P-Thr at position 701 by tryptic fingerprint analysis.

We have mentioned previously the presence of a 91K phosphorylated species (form Z) of large T in both SV40-transformed and SV40-infected permissive and nonpermissive cells (63). In this study we found that mild V8 proteolysis of these molecules yields a typical 34K fragment as well as a normal C-terminal 58K fragment. This 34K fragment was found to be related to the normal N-terminal 31K fragment. It is possible that Z and Z' molecules are translated from low-abundance mRNA with an alternative splicing pattern. Another possibility is that these species represent large T antigens binding to (a class of) RNA molecules or fragments thereof (26). Alternatively, Z molecules could be polyADPribosylated large T molecules. This type of modification has recently been reported to occur in a minor fraction of the large T molecules found in lytically infected cells (20). Up to now, there is no unequivocal evidence available to support the alternative splice hy-

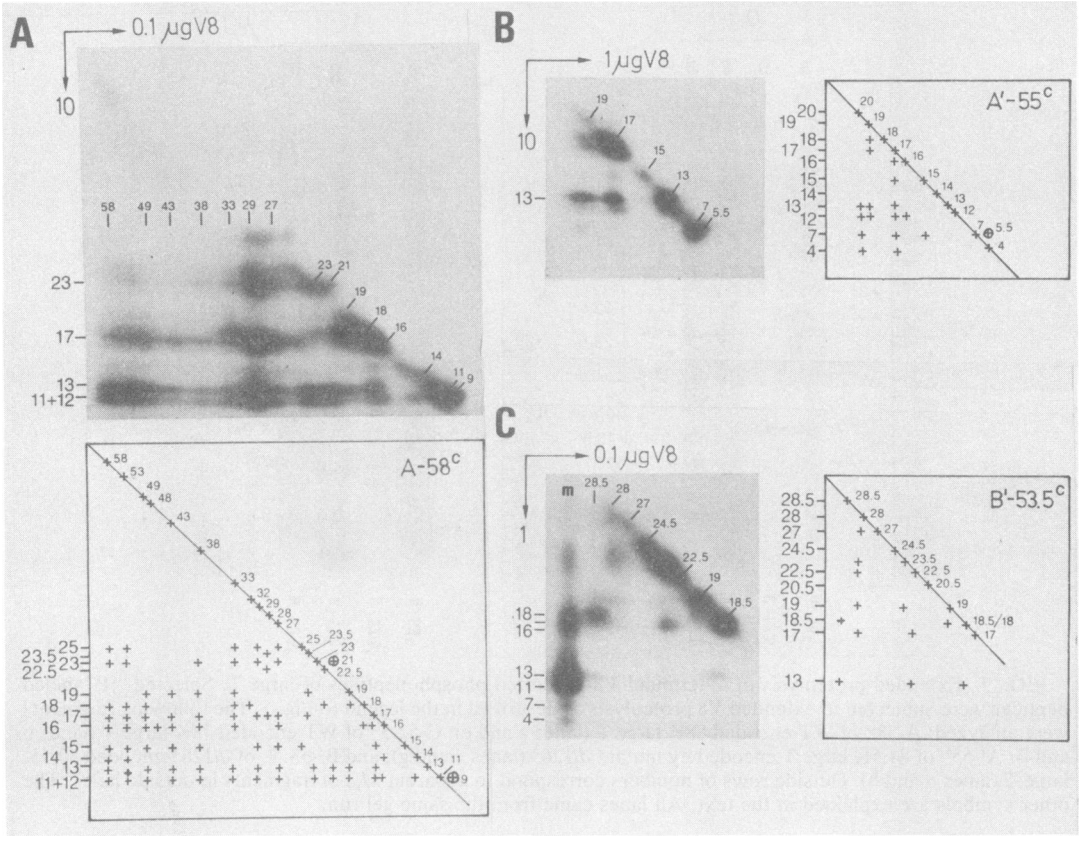


FIG. 10. Two-dimensional proteolysis mapping of C-terminal V8-generated phosphopeptides of large T. $^{32}P$-labeled peptides of large T were obtained by proteolysis with low amounts of V8 protease as described in the legend to Fig. 1. Extended V8 proteolysis of those peptides was performed as described in the legend to Fig. 9. Finally, strips of the latter gel were subjected to proteolysis with still higher amounts of V8 protease in a second dimension as described in the text. Each autoradiogram is clarified by a corresponding scheme. Horizontal and vertical arrows show directions of the first and second dimensions, respectively. The protease concentration used in each dimension is also indicated. The following phosphopeptides were analyzed: A-58$^C$ of WT-encoded 88K large T (A), A'-55$^C$ of 84.5K large T encoded by the mutant *dl*1263 (B), and B'-53.5$^C$ of *dl*1263-encoded 80.5K large T (C). The letter m in (C) indicates a one-dimensional reference digest of the A-58$^C$ peptide with 1 µg of V8 protease.

pothesis or the peculiar RNA-binding class, but our recent unpublished data do suggest that the Z molecules are polyADPribosylated to a greater extent [based on $^{32}P_i$ incorporation in 2'-(5"-phosphoribosyl)-5'-AMP] than A molecules. We have also found that the Z and A molecules bind to native DNA-cellulose with equal efficiency, and that both forms are complexed to host-coded 53K phosphoproteins in SV40-transformed cells (unpublished data).

Recently, information has become available on the minimum structural requirements of the target peptides for four different types of protein kinase. We therefore analyzed the amino acid sequence of large T for these putative substrate sequences (Table 3). Basically, cyclic AMP- and cyclic GMP-dependent protein kinases recognize the sequences Lys-Arg-X-(X)-Ser/Thr and Arg-Arg-X-(X)-Ser/Thr (16, 27). Ser residues are generally modified, but Thr has also been reported to be phosphorylated by these enzymes (1, 7, 35). There is only one sequence of this type in large T: Lys-Arg-Val-Asp-Ser$^{352}$. In addition, large T contains the sequence Lys-Gln-Val-Ser$^{269}$-tryptophan (Trp)-Lys, which is very similar to the sequence Lys-Gln-Ile-Ser-Val-Arg, known to be phosphorylated by the $Ca^{2+}$-dependent phosphorylase kinase (58). Large T also contains multiple sequences of the type Ser/Thr-X-Asp/Glu, recognized by the so-called cyclic AMP-independent casein kinase TS (4, 24, 31, 37). But the sequences typical for the so-called

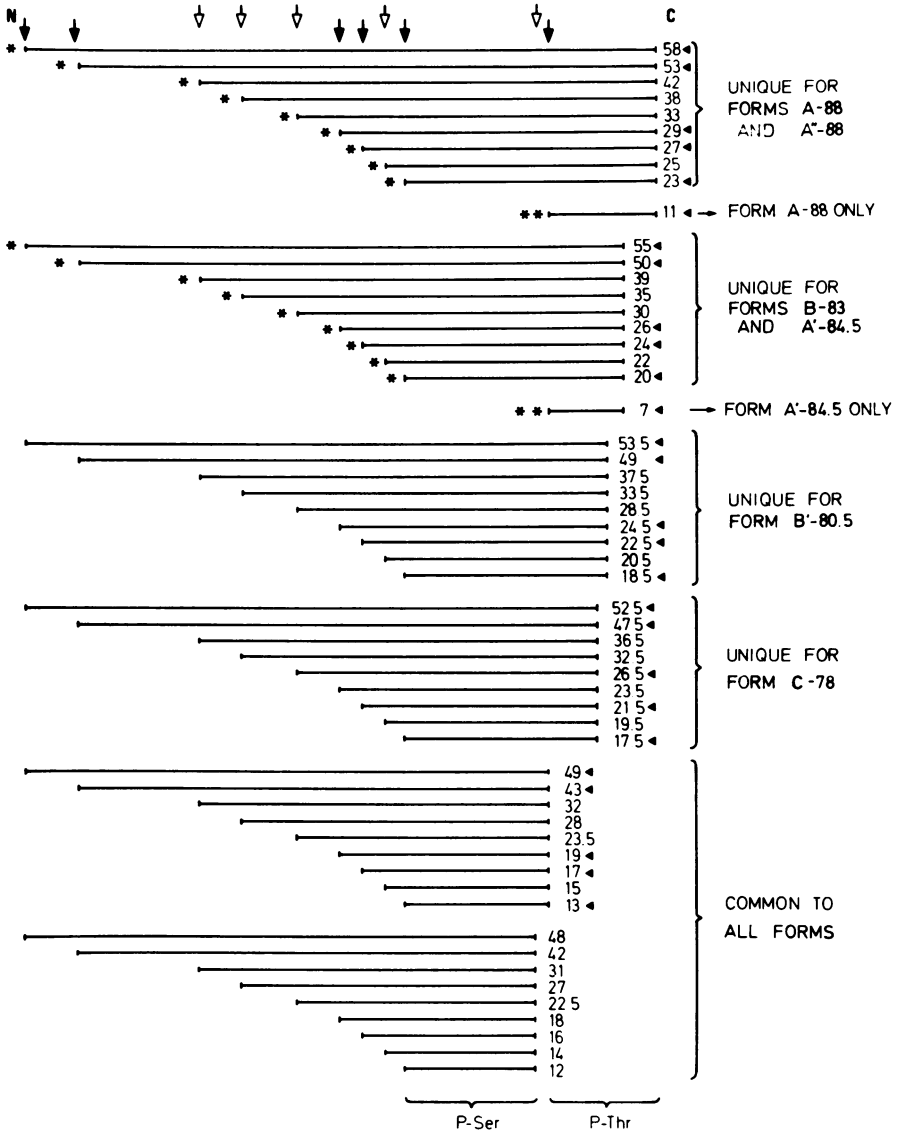


FIG. 11. Best-fitting model of location of detectable $^{32}$P-labeled proteolytic fragments in the C-terminal parts of different large T forms. All peptides were generated by *S. aureus* V8 digestion as illustrated in Fig. 8–10. The peptides have been ordered in this scheme according to the data discussed in the text. N and C indicate the N- and C-terminal ends of the peptides. Major peptides are indicated by horizontal arrowheads. Arrows with black heads indicate sites that are cleaved by V8, presumably by preference; arrows with white heads indicate minor cleavage sites. Peptides of A-88 and A'-84.5 large T forms which should contain both P-Ser and P-Thr residues are indicated by single asterisks. However, it should be noted that the same sets of peptides contain only P-Ser in the case of A"-88 and B-83 large T forms. Peptides A-11$^{C}$ and A'7$^{C}$, containing phosphate solely at position Thr$^{701}$, are indicated by double asterisks. All other peptides contain solely P-Ser residues. In summary, the maximum length of two phosphorylated domains was derived from these data, as shown at the bottom of this scheme. The first domain contains solely P-Ser residues. The second contains a single P-Thr residue at position 701.

cyclic AMP-independent casein kinase S are not present. The latter enzyme is a secondary kinase which phosphorylates Ser residues situated C-terminal to very acidic clusters, including several P-Ser residues (36). Except for the latter enzyme, protein kinases also generally require a β-turn as the secondary structure of their target sequence (36, 37, 44, 52). Indeed, β-turns are

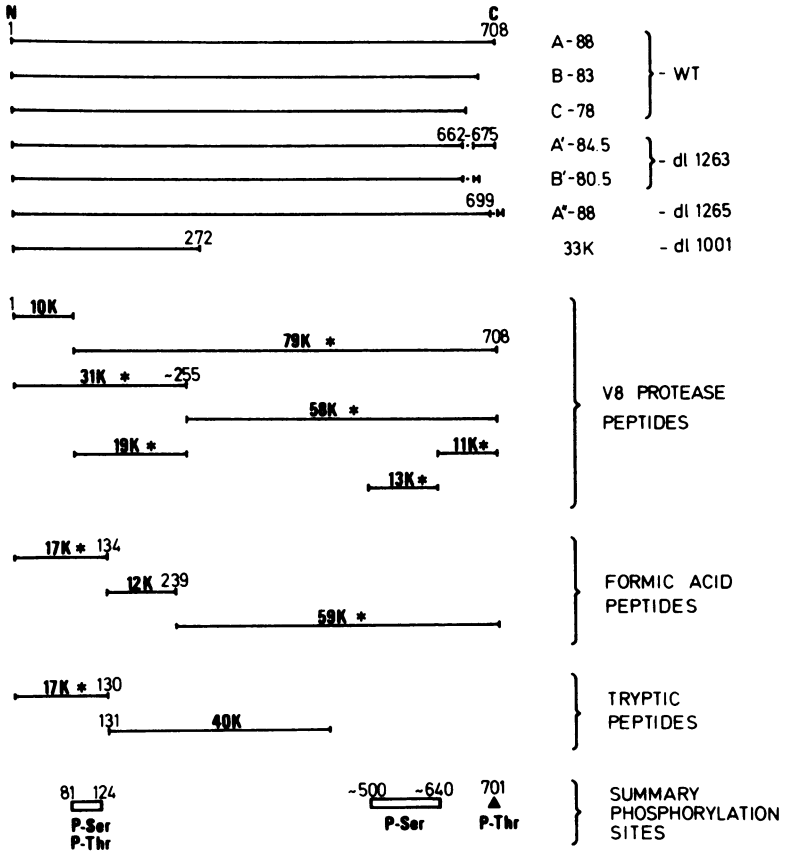


FIG. 12. Schematic representation of the presumptive phosphorylation sites in SV40 large T antigen. At the top are shown the different large T forms as described in this and a previous report (63). N and C indicate the N- and C-terminals, respectively. Numbers refer to positions in the amino acid sequence (60, 62). Below are shown major peptides generated by partial V8 proteolysis, by limited formic acid hydrolysis, or by partial tryptic proteolysis. The latter data were taken from Schwyzer et al. (49). Phosphorylated peptides are marked with asterisks. Shown at the bottom are two domains that carry phosphorylation sites and the P-Thr site at position 701. Further details are given in the text.

located predominantly at the periphery of the molecule, where they are ideally situated for recognition and modification by the kinase. Casein kinase TS sites also require that the β-turn forming tetrapeptides include both the target residue and the acidic determinant at the $n + 2$ position (36). After selection for this whole set of criteria, only the following sequences in large T remain as possible targets for the four types of kinases mentioned: $Thr^{88}$-Asp-Glu, $Ser^{111}$-$Ser^{112}$-Asp-Asp, $Thr^{117}$-Ala-Asp, $Ser^{632}$-Asp-Asp, and $Ser^{639}$-Gln-Glu (Table 3). It so happens that all six of these sites lie within the physical limits described above (Fig. 12). However, until the phosphorylation sites have been mapped unambiguously by phosphoamino acid sequencing, the possibility that large T contains additional substrate sequences cannot be excluded. Several other phosphoproteins have been found to contain such unclassified sequences, which are modified by the same or different protein kinases (27, 52, and references therein). One of these kinases undoubtedly phosphorylates $Thr^{701}$ (48), which does not fit in with the previous scheme except for the presence of a β-turn.

One can attempt to correlate phosphorylation of discrete sites in large T with any of its known physiological functions. Phosphorylation of the N-terminal part of large T could be involved in its DNA-binding capability. Mutant *dl*1001 is known to code for a 33K protein which contains only the first 272 amino acids of large T (47, 64). This protein still binds to double-stranded DNA (47) and contains both P-Ser and P-Thr residues (63). Furthermore, the adenovirus 2-SV40 hybrid $Ad2^{+}D2$ synthesizes a T-antigen-related

TABLE 3. Possible target sequences for different types of protein kinases in large T antigen

| Sequence[a] | Protein kinase[b] | $p_t$ ($\times 10^4$)[c] | $<P_t>$[c] | $<P_\alpha>$[c] | $<P_\beta>$[c] | Predicted as β-turn[d] | Presumed target sequence[e] |
|---|---|---|---|---|---|---|---|
| -[Gly-THR$^{88}$-Asp-*Glu*]- | TS | 1.26 | 1.18 | 0.98 | 0.71 | Yes | Yes |
| -[Cys-SER$^{106}$-Glu-*Glu*]- | TS | 1.02 | 1.02 | 1.12 | 0.67 | No | No |
| -[SER$^{111}$-SER$^{112}$-*Asp-Asp*]- | TS | 2.42 | 1.44 | 0.89 | 0.64 | Yes | Yes-Yes |
| -[THR$^{117}$-Ala-*Asp*-Ser]- | TS | 1.24 | 1.13 | 1.01 | 0.83 | Yes | Yes |
| -Lys-[Gln-Val-SER$^{269}$-Trp]-Lys-Leu- | $Ca^{2+}$ | 0.74 | 0.97 | 1.01 | 1.23 | No | No |
| -[Tyr-SER$^{298}$-Phe-*Glu*]- | TS | 0.47 | 0.98 | 1.02 | 0.99 | No | No |
| -*Lys-Arg*-Val-[Asp-SER$^{352}$-Leu-Gln]- | cAMP | 0.72 | 1.11 | 1.02 | 0.92 | No | No[f] |
| -[SER$^{381}$-Ala-*Asp*-Ile]- | TS | 0.91 | 1.00 | 1.07 | 0.93 | No | No |
| -[Asn-SER$^{632}$-Asp-*Asp*]- | TS | 3.24 | 1.48 | 0.86 | 0.68 | Yes | Yes |
| -[SER$^{639}$-Gln-*Glu*-Asn]- | TS | 0.82 | 1.18 | 1.01 | 0.78 | Yes | Yes |
| -[Pro-Pro-THR$^{701}$-Pro]- | ? | 1.36 | 1.38 | 0.63 | 0.71 | Yes | Yes[g] |

[a] The presumptive or actual target residues are in capitals and are numbered according to their position in large T (60, 62). Essential residues in the target sequences are italicized. Brackets enclose either the predicted β-turns or the tetrapeptides with the highest $p_t$ values. In the case of a casein TS substrate, these tetrapeptides also had to contain both the target residue and the acidic residue at position $n + 2$.

[b] TS, Casein kinase of type TS; $Ca^{2+}$, $Ca^{2+}$-dependent phosphorylase kinase; cAMP, cAMP-dependent protein kinase; ?, so far unclassified kinase.

[c] Calculated as described by Chou and Fasman (5). $P_t$, Probability of β-turn occurrence; $P_t$, conformational parameters of β-turns; $P_\alpha$, α-helices; β, β-sheets.

[d] Possible β-turns are tetrapeptides with $p_t > 0.75 \times 10^{-4}$ as well as $< P_t > > 1.00$ and $<P_\alpha> < <P_t> > <P_\beta>$ (5).

[e] In addition, the sequences -THR$^{164}$-Lys-*Glu*, -THR$^{237}$-Arg-*Asp*, -THR$^{356}$-Arg-*Glu*, and -SER$^{482}$-Arg-*Asp* in large T antigen are presumably not phosphorylated by casein kinase TS since they contain a basic residue at position $n + 1$ (44).

[f] Besides the absence of a β-turn, this is not an ideal phosphorylation site since the target residue is not preceded by a neutral hydrophobic residue (27).

[g] So far only, this sequence has been confirmed to be phosphorylated in vivo (48).

protein which lacks the first 82 amino acids of normal large T but still binds very efficiently to DNA (59). From recent results on the DNA-binding properties of T-antigen-related polypeptides coded for by the adenovirus 2-SV40 hybrid Ad2$^+$ND4, it was concluded that sequences in the region between 0.54 and 0.50 map units are required (46). Thus, the DNA-binding region of large T resides between approximate positions 82 and 140 and could, for instance, involve the basic cluster Lys$^{127}$-Lys-Lys-Arg-Lys$^{131}$ (64). The latter cluster is also present in the structural proteins VP2 and VP3, where it could play a similar role (68). In a neighboring or partly overlapping region between genome map coordinates 0.43 and 0.50, mutations were also mapped which result in a large T antigen binding more efficiently to a mutated origin of replication (51). Thus, phosphorylation of sites in this region can be expected to modulate DNA binding and can, therefore, have a role in regulating both viral replication and transcription (59). The same region of large T has also been found recently to be involved in the induction of cellular DNA synthesis (18).

The C-terminal P-Ser site could be situated in a region of large T which is largely present also in the 28K protein of the adenovirus-SV40 hybrid Ad2$^+$ND1. This virus contains the early SV40 DNA segment that extends from 0.17 to 0.29 map units (40). The 28K protein has been reported to function as a tumor-specific transplantation antigen (25) and is detectable at the cell surface (13). It also contains so-called U-antigenicity, which is associated with the nuclear envelope (12, 30). It may possibly be involved in viral DNA replication since, in contrast to cellular DNA induction, viral DNA replication does require the C-terminal end of large T (41). In this same region, there is also located a peculiar deletion which results in a phenotype which is heat sensitive for lytic growth but cold sensitive for transformation (45). Beyond doubt, the generation by site-directed mutagenesis of mutants lacking phosphorylation sites of large T will clarify this picture. Finally, it was mentioned previously (48, 63) that mutant *dl*1265 lacks the C-terminal P-Thr$^{701}$ site. This mutant is fully viable, although its adenovirus helper function is much impaired (8), suggesting an effect of either this segment or its phosphorylation on the latter activity.

## ACKNOWLEDGMENTS

We are grateful to C. Cole for providing us with SV40 mutants *dl*1263 and *dl*1265. We thank J. Van der Heyden for growth and titration of virus stocks, K. Van de Spiegle for expert help with tissue culture, and S. Kaplan for assistance with the preparation of the manuscript.

This investigation was supported by grants from the Fonds voor Geneeskundig Wetenschappelijk Onderzoek (no. 3.0033.80) and from the Kankerfonds of the Algemene Spaar- en Lijfrentekas of Belgium.

## LITERATURE CITED

1. **Aitken, A., T. Bilham, P. Cohen, D. Aswad, and P. Greengard.** 1981. A specific substrate from rabbit cerebellum for guanosine-3':5'-monophosphate dependent protein kinase. III. Amino acid sequences at the two phosphorylation sites. J. Biol. Chem. **256**:3501–3506.
2. **Amons, R., and P. I. Schrier.** 1981. Removal of sodium dodecyl sulfate from proteins and peptides by gel filtration. Anal. Biochem. **116**:439–443.
3. **Bradley, M. K., J. D. Griffin, and D. M. Livingston.** 1982. Relationship of oligomerization to enzymatic and DNA-binding properties of the SV40 large T antigen. Cell **28**:125–134.
4. **Brignon, G., B. Ribadeau Dumas, J.-C. Mercier, and J.-P. Pelissier.** 1977. Complete amino acid sequence of bovine $\alpha_{S2}$-casein. FEBS Lett. **76**:274–279.
5. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
6. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1102–1106.
7. **Cohen, P., D. B. Rylatt, and G. A. Nimmo.** 1977. The hormonal control of glycogen metabolism: the amino acid sequence at the phosphorylation site of protein phosphatase inhibitor-1. FEBS Lett. **76**:182–186.
8. **Cole, C. N., L. V. Crawford, and P. Berg.** 1979. Simian virus 40 mutants with deletions at the 3' end of the early region are defective in adenovirus helper function. J. Virol. **30**:683–691.
9. **Cole, C. N., T. Landers, S. P. 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.
10. **Crawford, L. V., and P. Z. O'Farrell.** 1979. Effect of alkylation on the physical properties of simian virus 40 T-antigen species. J. Virol. **29**:587–596.
11. **Denhardt, D. T., and L. V. Crawford.** 1980. Simian virus 40 T-antigen: identification of tryptic peptides in the C-terminal region and definition of the reading frame. J. Virol. **34**:315–329.
12. **Deppert, W.** 1979. Simian virus 40 T- and U-antigens: immunological characterization and localization in different nuclear subfractions of simian virus 40-transformed cells. J. Virol. **29**:576–586.
13. **Deppert, W., and R. Pates.** 1979. Cell surface location of simian virus 40-specific proteins on HeLa cells infected with adenovirus type 2-simian virus 40 hybrid viruses $Ad2^+ND1$ and $Ad2^+ND2$. J. Virol. **31**:522–536.
14. **Edwards, C. A. F., G. Khoury, and R. G. Martin.** 1979. Phosphorylation of T-antigen and control of T-antigen expression in cells transformed by wild-type and *tsA* mutants of simian virus 40. J. Virol. **29**:753–762.
15. **Fanning, E., B. Nowak, and C. Burger.** 1981. Detection and characterization of multiple forms of simian virus 40 large T antigen. J. Virol. **37**:92–102.
16. **Feramisco, J. R., D. B. Glass, and E. G. Krebs.** 1980. Optimal spatial requirements for the location of basic residues in peptide substrates for the cyclic AMP-dependent protein kinase. J. Biol. Chem. **255**:4240–4245.
17. **Feunteun, J., G. Carmichael, J. C. Nicolas, and M. Kress.** 1981. Mutant carrying deletions in the two simian virus 40 early genes. J. Virol. **40**:625–634.
18. **Galanti, N., G. J. Jonak, K. J. Soprano, J. Floros, L. Kaczmarek, S. Weissman, V. B. Reddy, S. M. Tilghman, and R. Baserga.** 1981. Characterization and biological activity of cloned simian virus 40 DNA fragments. J. Biol. Chem. **256**:6469–6474.
19. **Gidoni, D., A. Scheller, B. Barnet, P. Hantzopoulos, M. Oren, and C. Prives.** 1982. Different forms of simian virus 40 large tumor antigen varying in their affinities for DNA. J. Virol. **42**:456–466.
20. **Goldman, N., M. Brown, and G. Khoury.** 1981. Modification of SV40 T antigen by polyADP-ribosylation. Cell **24**:567–572.
21. **Greenspan, D. S., and R. B. Carroll.** 1979. Simian virus 40 large T antigen isoelectric focuses as multiple species with varying phosphate content. Virology **99**:413–416.
22. **Greenspan, D. S., and R. B. Carroll.** 1981. Complex of simian virus 40 large tumor antigen and 48,000 dalton host tumor antigen. Proc. Natl. Acad. Sci. U.S.A. **78**:105–109.
23. **Griffin, J. D., S. Light, and D. M. Livingston.** 1978. Measurements of the molecular size of the simian virus 40 large T antigen. J. Virol. **27**:218–226.
24. **Henderson, J. Y., A. J. G. Moir, L. A. Fothergill, and J. E. Fothergill.** 1981. Sequences of sixteen phosphoserine peptides from ovalbumins of eight species. Eur. J. Biochem. **114**:439–450.
25. **Jay, G., F. T. Jay, C. Chang, R. M. Friedman, and A. S. Levine.** 1978. Tumour-specific transplantation antigen: use of the Ad $2^+$ ND1 hybrid virus to identify the protein responsible for simian virus 40 tumor rejection and its genetic origin. Proc. Natl. Acad. Sci. U.S.A. **75**:3055–3059.
26. **Khandjian, E. W., M. Loche, J.-L. Darlix, R. Cramer, H. Türler, and R. Weil.** 1982. Simian virus 40 large tumor antigen: a "RNA binding protein"? Proc. Natl. Acad. Sci. U.S.A. **79**:1139–1143.
27. **Krebs, E. G., and J. A. Beavo.** 1979. Phosphorylation-dephosphorylation of enzymes. Annu. Rev. Biochem. **48**:923–959.
28. **Lam, K. S., and C. B. Kasper.** 1980. Sequence homology analysis of a heterogenous protein population by chemical and enzymatic digestion using a two-dimensional sodium dodecyl sulfate-polyacrylamide gel system. Anal. Biochem. **108**:220–226.
29. **Landon, M.** 1977. Cleavage at aspartyl-prolyl bonds. Methods Enzymol. **47**:145–149.
30. **Lin, P. S., R. Schmidt-Ullrich, and D. F. H. Wallach.** 1977. Transformation by simian virus 40 induces virus-specific, related antigens in the surface membrane and nuclear envelope. Proc. Natl. Acad. Sci. U.S.A. **74**:2495–2499.
31. **Mamrack, M. D., M. O. J. Olson, and H. Busch.** 1979. Amino acid sequence and sites of phosphorylation in a highly acidic region of nucleolar nonhistone protein C23. Biochemistry **18**:3381–3386.
32. **Mann, K., and T. Hunter.** 1980. Phosphorylation of SV40 large T antigen in SV40 nucleoprotein complexes. Virology **107**:526–532.
33. **McCormick, F., F. Chaudry, R. Harvey, R. Smith, P. W. J. Rigby, E. Paucha, and A. E. Smith.** 1980. T antigens of SV40-transformed cells. Cold Spring Harbor Symp. Quant. Biol. **44**:171–178.
34. **McCormick, F., and E. Harlow.** 1980. Association of a murine 53,000-dalton phosphoprotein with simian virus 40 large-T antigen in transformed cells. J. Virol. **34**:213–224.
35. **Meggio, F., G. Chessa, G. Borin, L. A. Pinna, and F. Marchiori.** 1981. Synthetic fragments of protamines as model substrates for rat liver cyclic-AMP-dependent protein kinase. Biochim. Biophys. Acta **662**:94–101.
36. **Meggio, F., A. Donella-Deana, and L. A. Pinna.** 1979. Studies on the structural requirements of a microsomal cAMP-independent protein kinase. FEBS Lett. **106**:76–80.

37. **Meggio, F., A. Donella-Deana, and L. A. Pinna.** 1981. A study with model substrates of the structure of the sites phosphorylated by rat casein kinase TS. Biochim. Biophys. Acta. **662**:1–7.
38. **Mellor, A., and A. E. Smith.** 1978. Characterization of the amino-terminal tryptic peptide of simian virus 40 small-t and large-T antigens. J. Virol. **28**:992–996.
39. **Montenarh, M., and R. Henning.** 1980. Simian virus 40 T-antigen phosphorylation is variable. FEBS Lett. **114**:107–110.
40. **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 chromosome. J. Virol. **12**:653–658.
41. **Mueller, C., A. Graessmann, and M. Graessmann.** 1978. Mapping of early SV40-specific functions by microinjection of different early viral DNA fragments. Cell **15**:579–585.
42. **Palme, K., and R. Henning.** 1980. Charge isomers of simian virus 40 T-antigen. FEBS Lett. **118**:229–232.
43. **Paucha, E., R. Harvey, and A. E. Smith.** 1978. Cell-free synthesis of simian virus 40 T-antigens. J. Virol. **28**:154–170.
44. **Pinna, L. A., A. Donella-Deana, and F. Meggio.** 1979. Structural features determining the site specificity of a rate liver cAMP-independent protein kinase. Biochem. Biophys. Res. Commun. **87**:114–120.
45. **Pintel, D., N. Bouck, and G. di Mayorca.** 1981. Separation of lytic and transforming functions of the simian virus 40 *A* region: two mutants which are temperature sensitive for lytic functions have opposite effects on transformation. J. Virol. **38**:518–528.
46. **Prives, C., B. Barnet, A. Scheller, G. Khoury, and G. Jay.** 1982. Discrete regions of simian virus 40 large T antigen are required for nonspecific and viral origin-specific DNA binding. J. Virol. **43**:73–82.
47. **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.
48. **Scheidtmann, K.-H., A. Kaiser, A. Carbone, and G. Walter.** 1981. Phosphorylation of threonine in the proline-rich carboxy-terminal region of simian virus large T antigen. J. Virol. **38**:59–69.
49. **Schwyzer, M., R. Weil, G. Frank, and H. Zuber.** 1980. Amino acid sequence analysis of fragments generated by partial proteolysis from large simian virus 40 tumor antigen. J. Biol. Chem. **255**:5627–5634.
50. **Shaw, S. B., and P. Tegtmeyer.** 1981. Binding of dephosphorylated A protein to SV40 DNA. Virology **115**:88–96.
51. **Shortle, D. R., R. F. Margolskee, and D. Nathans.** 1979. Mutational analysis of the simian virus 40 replicon: pseudorevertants of mutants with a defective replication origin. Proc. Natl. Acad. Sci. U.S.A. **76**:6128–6131.
52. **Small, D., P. Y. Chou, and G. D. Fasman.** 1977. Occurrence of phosphorylated residues in predicted β-turns: implications for β-turn participation in control mechanisms. Biochem. Biophys. Res. Commun. **79**:341–346.
53. **Smith, A. E., R. Smith, and E. Paucha.** 1978. Extraction and fingerprint analysis of simian virus 40 large and small T-antigens. J. Virol. **28**:140–153.
54. **Smith, A. E., R. Smith, and E. Paucha.** 1979. Characterization of different tumor antigens present in cells transformed by simian virus 40. Cell **18**:335–346.
55. **Spangler, G. J., J. D. Griffin, H. Rubin, and D. M. Livingston.** 1980. Identification and initial characterization of a new low-molecular-weight virus-encoded T antigen in a line of simian virus 40-transformed cells. J. Virol. **36**:488–498.
56. **Stitt, D. T., R. B. Carroll, J. A. Melero, and W. F. Mangel.** 1981. Analysis of the 84K, 55K, and 48K proteins immunoprecipitable by SV40 T antibody from SV40-infected and transformed cells by tryptic peptide mapping on cation-exchange columns. Virology **111**:283–288.
57. **Tegtmeyer, P., K. Rundell, and J. K. Collins.** 1977. Modification of simian virus 40 protein A. J. Virol. **21**:647–657.
58. **Tessmer, G. W., J. R. Skuster, L. B. Tabatabai, and D. J. Graves.** 1977. Studies on the specificity of phosphorylase kinase using peptide substrates. J. Biol. Chem. **252**:5666–5671.
59. **Tjian, R.** 1981. Regulation of viral transcription and DNA replication by the SV40 large T antigen. Curr. Top. Microbiol. Immunol. **93**:5–24.
60. **Tooze, J. (ed).** 1980. The molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
61. **Van Heuverswyn, H., C. Cole, P. Berg, and W. Fiers.** 1979. Nucleotide sequence analysis of two simian virus 40 mutants with deletions in the region coding for the carboxyl terminus of the T antigen. J. Virol. **30**:936–941.
62. **Van Heuverswyn, H., A. Van De Voorde, J. Van Herreweghe, G. Volckaert, P. De Winne, and W. Fiers.** 1980. Nucleotide sequence of the simian virus 40 Hind II and III restriction fragment B (sixth part of the T antigen gene) and codon usage. Eur. J. Biochem. **106**:199–209.
63. **Van Roy, F., L. Fransen, and W. Fiers.** 1981. Phosphorylation patterns of tumor antigens in cells lytically infected or transformed by simian virus 40. J. Virol. **40**:28–44.
64. **Volckaert, G., A. Van de Voorde, and W. Fiers.** 1980. Nucleotide sequence of the simian virus 40 Hind II + III restriction fragment A (second part of the T antigen gene). Eur. J. Biochem. **106**:169–177.
65. **Walter, G., and P. J. Flory, Jr.** 1980. Phosphorylation of SV40 large T antigen. Cold Spring Harbor Symp. Quant. Biol. **44**:165–169.
66. **Walter, G., K.-H. Scheidtmann, A. Carbone, A. P. Laudano, and R. F. Doolittle.** 1980. Antibodies specific for the carboxy- and amino-terminal regions of simian virus 40 large tumor antigen. Proc. Natl. Acad. Sci. U.S.A. **77**:5197–5200.
67. **Yang, Y.-C., P. Hearing, and K. Rundell.** 1979. Cellular proteins associated with simian virus 40 early gene products in newly infected cells. J. Virol. **32**:147–154.
68. **Ysebaert, M., A. Van de Voorde, and W. Fiers.** 1978. Nucleotide sequence of the simian virus 40 Hind II and III restriction fragment D and the total amino acid sequence of the late proteins VP2 and VP3. Eur. J. Biochem. **91**:431–439.