JOURNAL OF VIROLOGY, Oct. 1982, p. 116–133
0022-538X/82/100116-18$02.00/0

Vol. 44, No. 1

# Simian Virus 40 Large T Antigen Is Phosphorylated at Multiple Sites Clustered in Two Separate Regions

KARL-HEINZ SCHEIDTMANN,* BIRGIT ECHLE, AND GERNOT WALTER

*Institut fuer Immunbiologie, Universitaet Freiburg, D-7800 Freiburg, Federal Republic of Germany*

Received 22 March 1982/Accepted 11 June 1982

The phosphorylation sites of simian virus 40 large T antigen were determined within the primary structure of the molecule. Exhaustive digestion of $^{32}P$-labeled large T antigen with trypsin generated six major phosphopeptides which could be separated in a newly developed isobutyric acid-containing chromatography system. By partial tryptic digestion, large T antigen was cleaved into an amino-terminal fragment of 17,000 daltons and overlapping fragments from the carboxy-terminal region ranging in size between 71,000 and 13,000 daltons. The location of the phosphopeptides was then determined by fingerprint analyses of individual fragments. Their physical properties were analyzed by sizing on polyacrylamide gels and by sequential digestion and peptide mapping; their amino acid composition was determined by differential labeling with various amino acids. The amino-terminal 17,000-dalton fragment gave rise to only one phosphopeptide (phosphopeptide 3) that contained half of the phosphate label incorporated into large T antigen. It contained phosphoserine and phosphothreonine sites, all of which were clustered within a small segment between $Cys_{105}$ and $Lys_{127}$. This segment contained five serines and two threonines. Among these, $Ser_{106}$, $Ser_{123}$, and $Thr_{124}$ were identified as phosphorylated residues; in addition, either one or both of $Ser_{111}$ and $Ser_{112}$ were phosphorylated. The neighboring residues, $Ser_{123}$ and $Thr_{124}$, were found in three different phosphorylation states in that either $Ser_{123}$ or $Thr_{124}$ or both were phosphorylated. Phosphopeptides 1, 2, 4, 5, and 6 were all derived from a single fragment extending 26,000 daltons upstream from the carboxy terminus of large T antigen. Phosphopeptide 6 was identical with the previously determined phosphothreonine peptide phosphorylated at $Thr_{701}$. Phosphopeptides 1, 2, 4, and 5 contained only serine-bound phosphate. Phosphopeptides 1, 2, and 4 represented overlapping peptides, all of which were phosphorylated at $Ser_{639}$ located next to a cluster of six acidic residues. In phosphopeptide 5, a large peptide ranging from $Asn_{653}$ to $Arg_{691}$, at least two of seven serines were phosphorylated. Thus, large T antigen contains at least eight phosphorylation sites. Their clustering within two separate regions might correlate with structural and functional domains of this protein.

The simian virus 40 (SV40) large T antigen (large T) is a phosphoprotein of 82,000 daltons (82K) (36, 46). It plays a central role in the growth cycle of the virus as well as in virus-induced malignant transformation. During lytic infection large T initiates replication and regulates transcription of the viral genome (reviewed in reference 49). In vitro, large T binds to DNA, specifically to the regulatory region on the SV40 genome that includes the origin of DNA replication (19, 40, 45, 47); moreover, it is associated with an ATPase and a protein kinase activity (1, 17, 48).

How these different functions are exerted by a single protein is not understood, but it is possible that large T can assume different functional forms which are interconvertible by phosphorylation. Compatible with this possibility are the following findings: large T is phosphorylated at multiple sites (52) in a reversible fashion (9, 38); subclasses of large T can be separated by isoelectric focusing (15); and correlations exist between the degree of phosphorylation of large T on one hand and its degree of oligomerization (10, 16) or its affinity for DNA (31) on the other. However, the functional significance of these quantitative differences in the phosphorylation state is not yet clear. Montenarh and Henning reported that highly phosphorylated large T exhibits the strongest affinity for (calf thymus) DNA (31), whereas Shaw and Tegtmeyer found no difference between phosphorylated and dephosphorylated large T in its capacity to bind to the replication origin of the SV40 genome (41).

To assess a specific role of phosphorylation in the various functions and interactions of large T, detailed information about individual phosphorylation sites and their location within the polypeptide chain and their appearance in different subclasses of large T will be required.

A preliminary localization of the phosphorylation sites has been achieved by determining the distribution of phosphate in defined fragments of large T. Studies with the deletion mutant *dl*1001 (36) or with adenovirus-SV40 hybrid viruses (52) suggested that the majority of phosphorylation sites of large T might be located in an amino-terminal region encoded from 0.625 to 0.59 map units or from 0.54 to 0.44 map units (or both) on the SV40 genome. Schwyzer et al. (39) analyzed the phosphate content of specific fragments of large T which were generated by partial proteolysis. These authors found that the majority of phosphates were bound to an amino-terminal fragment of 17K extending to $Arg_{130}$, that a central region of 40K distal from $Arg_{130}$ was essentially free of phosphate, and that an additional phosphorylation site(s) must be located in a carboxy-terminal region of 30K. Similar results were obtained by other investigators (51).

We have previously localized the major phosphothreonine site in the carboxy-terminal tryptic peptide. A characterization of the phosphoserine peptides was not possible because we were not able to resolve them in the usual peptide mapping systems (37).

In the present paper we describe a newly developed chromatography system capable of separating highly polar and large peptides. Using this system we have identified six major phosphopeptides and analyzed their distribution within defined proteolytic fragments of large T (39). The phosphopeptides were further characterized by sizing on polyacrylamide gels, by sequential cleavage with proteases, and by differential labeling with various amino acids. At least three phosphorylated serines and one phosphothreonine were found in the amino-terminal region; all four were clustered between $Ser_{106}$ and $Thr_{124}$. The carboxy-terminal region contained at least three phosphoserine sites in addition to the previously identified phosphothreonine.

## MATERIALS AND METHODS

**Cell line and virus.** The TC7 subclone of the CV1 African green monkey kidney cell line was used in all experiments. The cells were grown in Dulbecco modified Eagle minimal essential medium supplemented with 5% fetal calf serum (Boehringer, Mannheim, Federal Republic of Germany). The large-plaque strain of SV40, originally obtained from P. Tegtmeyer, was propagated in TC7 cells. Virus stocks were kept at −60°C and had a titer of approximately $5 \times 10^7$ PFU/ml.

**Infection and labeling of cells.** Confluent TC7 cells growing on 9-cm plastic petri dishes were infected with SV40 at a multiplicity of 5 to 10 PFU/cell. Cells were routinely labeled between 40 and 50 h postinfection in 1.5 ml of Eagle medium with 5% fetal calf serum for 4 h. Labeling with carrier-free $^{32}P_i$ was carried out at 1 mCi per plate in phosphate-free Eagle medium with 5% fetal calf serum. $^3H$- or $^{14}C$-labeled essential amino acids were used at 1 or 0.125 mCi per plate, respectively, in Eagle medium with 5% fetal calf serum lacking the respective amino acid. Labeling with the nonessential amino acids was in 3 ml of complete Eagle medium with 5% fetal calf serum from 24 to 44 h postinfection. [$^{35}S$]methionine and [$^{35}S$]cysteine were used at 0.25 mCi per plate. All radiolabeled materials were obtained from Amersham International, Ltd., Amersham, United Kingdom.

**Isolation of large T.** Cytoplasmic extracts from infected and radiolabeled cells were prepared by low-salt extraction in 0.5% Nonidet P-40 as described previously (30, 37), but all Tris buffers were replaced by $NaPO_4$ buffers because 60 to 100% higher yields of phosphate label bound to large T were obtained with $NaPO_4$ buffers. Isolation of large T by immunoprecipitation and sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described previously (22, 27, 30).

After autoradiography of the unfixed dried gel large T was eluted in buffer containing 50 mM $NH_4HCO_3$, 0.1% SDS, and 5% 2-mercaptoethanol; it was precipitated with 20% trichloroacetic acid and oxidized with 3% performic acid (37).

**Partial digestion of large T immune complexes.** When large T-antibody complexes are treated with proteases under mild conditions, large T is cleaved only at preferential sites; some material may be released after cleavage, but most of the digestion products remain tightly associated with the immune complex (39). The original procedure described by Schwyzer et al. (39) was followed, except that fixed *Staphylococcus aureus* instead of protein A-Sepharose was used for immunoprecipitation. Briefly, large T was immunoprecipitated with hamster antitumor serum and *S. aureus* as described previously (30). The precipitate was washed twice in precipitation buffer (30), once in pH 9 lysis buffer (10 mM Tris hydrochloride [pH 9.0], 100 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 0.5 mM $MgCl_2$, 0.5% Nonidet P-40 [39]), suspended in 1/4 to 1/5 of the original volume in pH 9 lysis buffer, and incubated with tolylsulfonyl phenylalanyl chloromethyl ketone-trypsin at a concentration of 30 μg/ml at 0°C for 30 min. The immunoprecipitate was sedimented, and the supernatant containing released material was adjusted to 2% SDS–20 mM dithiothreitol and electrophoresed on 12.5% SDS-polyacrylamide gels. The immune complex was dissociated with SDS-PAGE sample buffer, and the digestion products were separated on SDS-polyacrylamide gels consisting of a 7-cm layer of 15% polyacrylamide and a 6-cm layer of 10% polyacrylamide.

Elution, precipitation with trichloroacetic acid, and oxidation were performed as described above for intact large T.

**Two-dimensional peptide analysis.** Digestion of large T or of fragments of large T with tolylsulfonyl phenylalanyl chloromethyl ketone-trypsin or chymotrypsin or pronase E was carried out in 0.5 ml of 50 mM

$NH_4HCO_3$ (pH 7.9) at an enzyme concentration of 20 μg/ml at 37°C two times for 5 h. All enzymes were preincubated at 37°C for 60 min to destroy contaminating phosphatase and protease activities. In sequential digests the enzyme used first was inactivated by boiling. The digestion products were separated by two-dimensional fingerprinting on thin-layer cellulose (TLC) plates as described by Gibson (13).

Different electrophoresis or chromatography buffers were used depending on whether separation of the phosphopeptides or of unphosphorylated peptides was desired.

Electrophoresis in the first dimension was carried out in 6% formic acid–1.25% acetic acid–0.25% (vol/vol) pyridine, pH 1.9 (buffer A) (37); in 5% butanol–2.5% acetic acid–2.5% (vol/vol) pyridine, pH 4.7 (buffer B) (30); or in 1% (wt/vol) $NH_4HCO_3$, pH 8.9 (buffer C) (52).

Ascending chromatography was performed in 39.3% butanol–30.4% pyridine–6% (vol/vol) acetic acid (buffer D) (30) or in 37.5% butanol–25% pyridine–7.5% (vol/vol) acetic acid (buffer E) (52). For the highly polar phosphopeptides a new chromatography system was developed, composed of isobutyric acid-pyridine-acetic acid-butanol-water (65:5:3:2:29) (buffer F). The positions of $^{35}S$- or $^{14}C$-labeled peptides were visualized by fluorography (3); $^{32}P$-labeled peptides were visualized by direct autoradiography with an intensifying screen. Comigration between $^{35}S$- or $^{14}C$-labeled and $^{32}P$-labeled peptides was assessed by mixing digests of either kind, peptide mapping, and autoradiography. To visualize the phosphopeptides alone, a fourfold layer of aluminium foil was placed between the TCL plate and X-ray film. $^{3}H$-labeled samples were mixed with $^{32}P$-labeled digests of large T and fingerprinted as described above; $^{32}P$-labeled peptides were identified by autoradiography, eluted from the cellulose, and analyzed for their $^{3}H$ content by liquid scintillation counting (37).

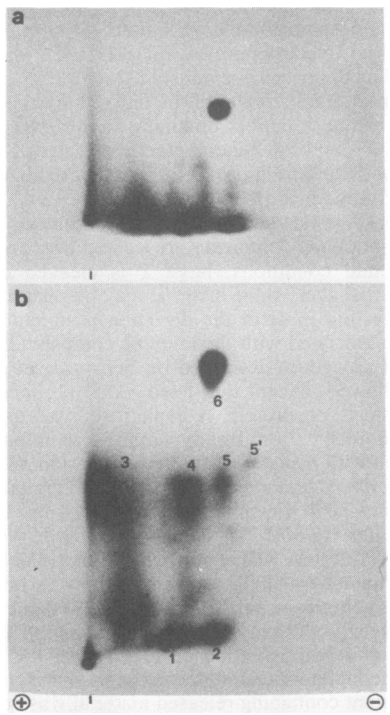


FIG. 1. Two-dimensional separation of tryptic phosphopeptides of large T. SV40-infected TC7 cells were labeled with $^{32}P_i$ from 44 to 48 h postinfection. Large T was purified by immunoprecipitation and SDS-PAGE and digested with tolylsulfonyl phenylalanyl chloromethyl ketone-trypsin as described in the text. The resulting peptides were separated on TLC plates by electrophoresis at pH 1.9 at 1,000 V for 35 min in the horizontal direction and by ascending chromatography in buffer E (a) or in buffer F (b) for 5 h. The origin is indicated by a vertical dash. The TLC plates were autoradiographed on Kodak X-Omat AR film with an intensifying screen.

**Determination of phosphoamino acids.** Acid hydrolysis of isolated phosphopeptides was performed in 5.6 N HCl at 110°C for 2 h in sealed glass tubes. After a 20-fold dilution in water, the acid was removed by lyophilization over an NaOH pellet. The hydrolysates were mixed with unlabeled phosphoserine and phosphothreonine, applied onto TLC plates, and electrophoresed in buffer A at 1,000 V for 70 min. The markers were visualized with ninhydrin (37).

**Size analysis of phosphopeptides by SDS-PAGE.** Proteolytic digests of $^{32}P$-labeled large T were lyophilized and dissolved in SDS-PAGE sample buffer and analyzed on linear gradient gels containing 10 to 30% polyacrylamide. The following size markers (and molecular weights) were used: $^{14}C$-labeled carbonic anhydrase (30,000), soy bean trypsin inhibitor (21,500), cytochrome *c* (12,500), aprotinin (6,500), and insulin β chain (3,400) and α chain (2,350). Molecular weight markers were purchased from Amersham.

## RESULTS

**Separation of the phosphopeptides of large T.** To study the phosphorylation of SV40 large T during productive infection, SV40-infected monkey cells were labeled with $^{32}P$ late during infection. Large T was isolated by immunoprecipitation and SDS-PAGE as described previously (30). After digestion of large T with trypsin, the phosphopeptides were separated by two-dimensional fingerprinting on TLC plates. In previously published work we were able to separate and analyze the major phosphothreonine-containing peptide, whereas the bulk of the phosphoserine-containing peptides could not be resolved due to their apparent heterogeneity and polarity (37) (Fig. 1a). Even sequential digestion with various proteases and peptide mapping in a variety of commonly used systems did not result in sufficient resolution of the phosphopeptides. The new chromatography system introduced in the present investigation is based on an isobutyric acid buffer (isobutyric acid–0.5 M $NH_3$, 5:3) originally used by Eckhart et al. to separate phosphothreonine and phosphotyrosine (8). When we used this buffer to

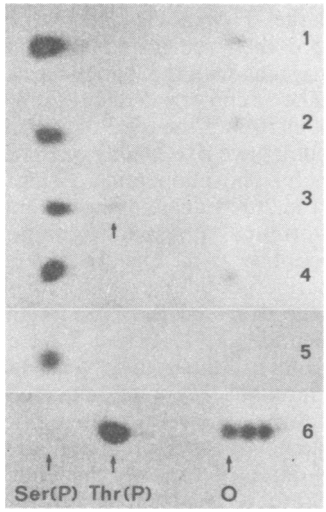


FIG. 2. Analysis of phosphoamino acids of isolated peptides. The phosphopeptides shown in Fig. 1b were isolated from TLC plates, hydrolyzed as described in the text, and electrophoresed at pH1.9 at 1,500 V for 60 min. The numbers correspond to individual phosphopeptides in Fig. 1b. O, Origin.

separate the phosphopeptides of large T, they migrated far up, some of them with the buffer front. Therefore, we tried various combinations of the isobutyric acid buffer with the chromatography buffer used in our previous studies (buffer E, see above). By slightly varying the ratios between polar and apolar constituents we observed great differences in the relative mobilities of different peptides. Thus, this chromatography system should be applicable to many specific separation problems. The buffer used in the following peptide mapping experiments was composed of isobutyric acid-pyridine-acetic acid-butanol-water (65:5:3:2:29) (buffer F).

When a tryptic digest of $^{32}P$-labeled large T was analyzed by electrophoresis at pH 1.9 and subsequent chromatography in buffer F, six major phosphopeptides could be separated designated phosphopeptides 1 through 6 (Fig. 1b). Phosphopeptides 3 and 4 appeared to be rather heterogeneous, suggesting that they were either incompletely digested or varied in their extent of phosphorylation. Phosphopeptide 3' was not reproducibly obtained and is probably a degradation product of phosphopeptide 3; phosphopeptides 5 and 5' are also related to each other (see below). Phosphopeptide 6 represents the carboxy-terminal phosphothreonine-containing peptide identified previously (37). To investigate the nature of the phosphate linkage in the individual peptides, they were isolated from the cellulose plates and analyzed by acid hydrolysis and subsequent electrophoresis. Phosphopeptides 1, 2, 4, and 5 contained only phosphoserine; phosphopeptide 3 (and 3') contained both phosphoserine and phosphothreonine in a ratio of 6:1; and as expected, phosphopeptide 6 contained only phosphothreonine (Fig. 2). The latter peptide appeared to be incompletely hydrolyzed since some material remained at the origin. Presumably, the proline bonds in this peptide are partially resistant to hydrolysis under the conditions employed.

**Mapping of phosphopeptides by partial proteolysis of large T.** A preliminary localization of phosphorylation sites of large T has been achieved by Schwyzer et al. (39) by analyzing partial cleavage products of large T for their phosphate content. Upon partial tryptic digestion, large T is cleaved into an amino-terminal fragment of 17K and several overlapping fragments of 71K, 67K, 58K, and 40K sharing $Lys_{131}$ as the amino-terminal residue and extending to various degrees toward the carboxy terminus. Among these fragments, only the 17K, 71K, 67K, and 58K fragments, but not the 40K fragment, contain phosphate (39). Thus, peptide analysis of the individual fragments should allow the alignment of the phosphopeptides with distinct regions of large T. With a slightly modified procedure for partial digestion (described above), fragments of 71K, 67K, 58K, 40K, 26K, 24K, 22K, and 17K were obtained. All of these contained phosphate, except for the 40K fragment (Fig. 3). This result is in agreement with the data published by Schwyzer et al. (39), except that these authors did not observe the 26K, 24K, and 22K fragments. The reason for this difference is not known. The alignment of the fragments within the large T polypeptide was determined by tryptic peptide analysis of [$^{35}S$]methionine-labeled fragments and comparison with the established arrangement of methionine-containing peptides (25, 30) (Fig. 4). The nomenclature is according to Mann et al. (30) and Linke et al. (25). Figure 4a shows a peptide map of complete large T. The 17K fragment (Fig. 4b) contained all peptides that are common between large T and small t (25). The 71K and 67K fragments (Fig. 4c) contained all peptides that have been ascribed to the unique portion of large T (25, 30). The 58K fragment (Fig. 4d) lacked only one peptide (methionine peptide 2), which was the most carboxy-terminal methionine peptide (see below). The 26K and 22K fragments (Fig. 4e) consisted of peptides deriving from the carboxy-terminal region (30, see below). Figure 4f shows a digest of the 22K fragment, which was fingerprinted by using the isobutyric acid-containing chromatography system to resolve peptide 2 (see below). The data demonstrated in Fig. 4b through d are consistent with the findings of Schwyzer et al. (39) that the 17K fragment is derived from the amino termi-

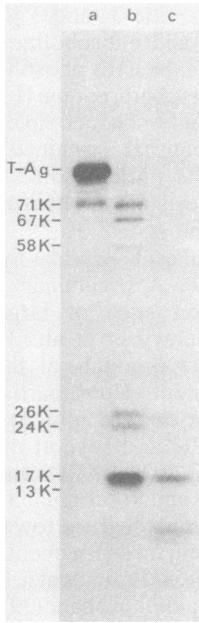


FIG. 3. Analysis of partial digestion products of $^{32}$P-labeled large T. Large T was extracted from SV40-infected cells immunoprecipitated and partially digested with trypsin as described in the text. The digestion products were analyzed on a gel consisting of a 7-cm layer of 15% polyacrylamide and a 6-cm layer of 10% polyacrylamide. Lanes: a, large T, untreated; b, fragments of large T eluted from the immune complex after digestion; c, fragments released during digestion.

nus of large T and that the 71K, 67K, and 58K fragments are derived from the unique part of large T, sharing a common amino terminus but differing at their carboxy termini. As will be shown below, the 26K and 22K fragments have the same carboxy termini as the 71K and 67K fragments, respectively. Since the difference in molecular weight between the 26K and 22K fragments is the same as that of the 71K and 67K fragments, we can assume that the former also have common amino termini. Although the extension of a given fragment cannot be deduced precisely from its apparent molecular weight, we assume that the 26K and 22K fragments extend to $Thr_{518}$ (see sequence in Fig. 9) because the sequence $\cdot$ Lys $\cdot$ Arg $\cdot$ $Thr_{518}$ $\cdot$ might be easily accessible during partial digestion with trypsin, and because the peptide $Thr_{518} \cdots Lys_{535}$ was the most amino-terminal methionine peptide obtained from the 26K and 22K fragments as identified by differential labeling. This peptide corresponds to methionine peptide 5 in Fig. 4e and f.

To determine the distribution of phosphate and the origin of the phosphopeptides, $^{32}$P-labeled large T was cleaved into fragments which were then redigested and analyzed by peptide mapping with the isobutyric acid system (Fig. 5). The amino-terminal 17K fragment contained about 50% of the total phosphate label of large T, but it gave rise to only one phosphopeptide, namely, phosphopeptide 3 (Fig. 5b). The 71K and the 26K fragments both contained the carboxy-terminal phosphothreonine peptide (phosphopeptide 6 in Fig. 5d and g) and all phosphoserine peptides except phosphopeptide 3. From the 67K and the 22K fragments only phosphopeptides 1, 2, 4, and 5 were obtained (Fig. 5e and h); thus, they were lacking the carboxy terminus. Peptide maps of the 24K fragment always showed traces of phosphopeptide 6 (data not shown), probably due to incomplete separation of the 26K and the 24K fragments. The 58K fragment (Fig. 5f) contained only phosphopeptides 1, 2, and 4, but traces of phosphopeptide 5 were sometimes observed. These results demonstrate that the phosphorylation sites of large T are located in two separate regions, an amino-terminal region of 17K and a region extending 26K upstream from the carboxy terminus.

To localize the phosphorylated residues more precisely within the amino acid sequence of large T, the phosphopeptides were further analyzed. Their physical properties were investigated by SDS-PAGE and peptide mapping of sequential digests. The amino acid composition of individual phosphopeptides was elucidated by differential labeling with various radiolabeled amino acids as described previously (37).

**Analysis of the amino terminal phosphopeptide 3.** Examination of the predicted amino acid sequence of the 17K fragment (11, 34) reveals that it contains three tryptic serine peptides. Two of these are relatively small and hydrophobic; they contain methionine and are probably identical with methionine peptides N and P or Q (Fig. 4b). The third peptide is a large peptide of 60 to 64 amino acids extending from $Tyr_{68}$ to a cluster of five basic residues around $Arg_{130}$. To determine the size of phosphopeptide 3, tryptic digests of $^{32}$P-labeled large T or of the 17K fragment were analyzed on 10 to 30% polyacrylamide gradient gels (Fig. 6). Digestion of large T revealed large phosphopeptides heterogeneously migrating with apparent molecular weights of 5K to 7K and smaller peptides migrating with the front (Fig. 6, lane b); digestion of the 17K fragment usually yielded only a large phosphopeptide of 6.5K, but sometimes small material which migrated with the front was also observed (Fig. 6, lane g). When phosphopeptide 3 was eluted from a TLC plate and analyzed on a polyacrylamide gel it migrated as a 6.5K peptide. These data suggested that phosphopeptide

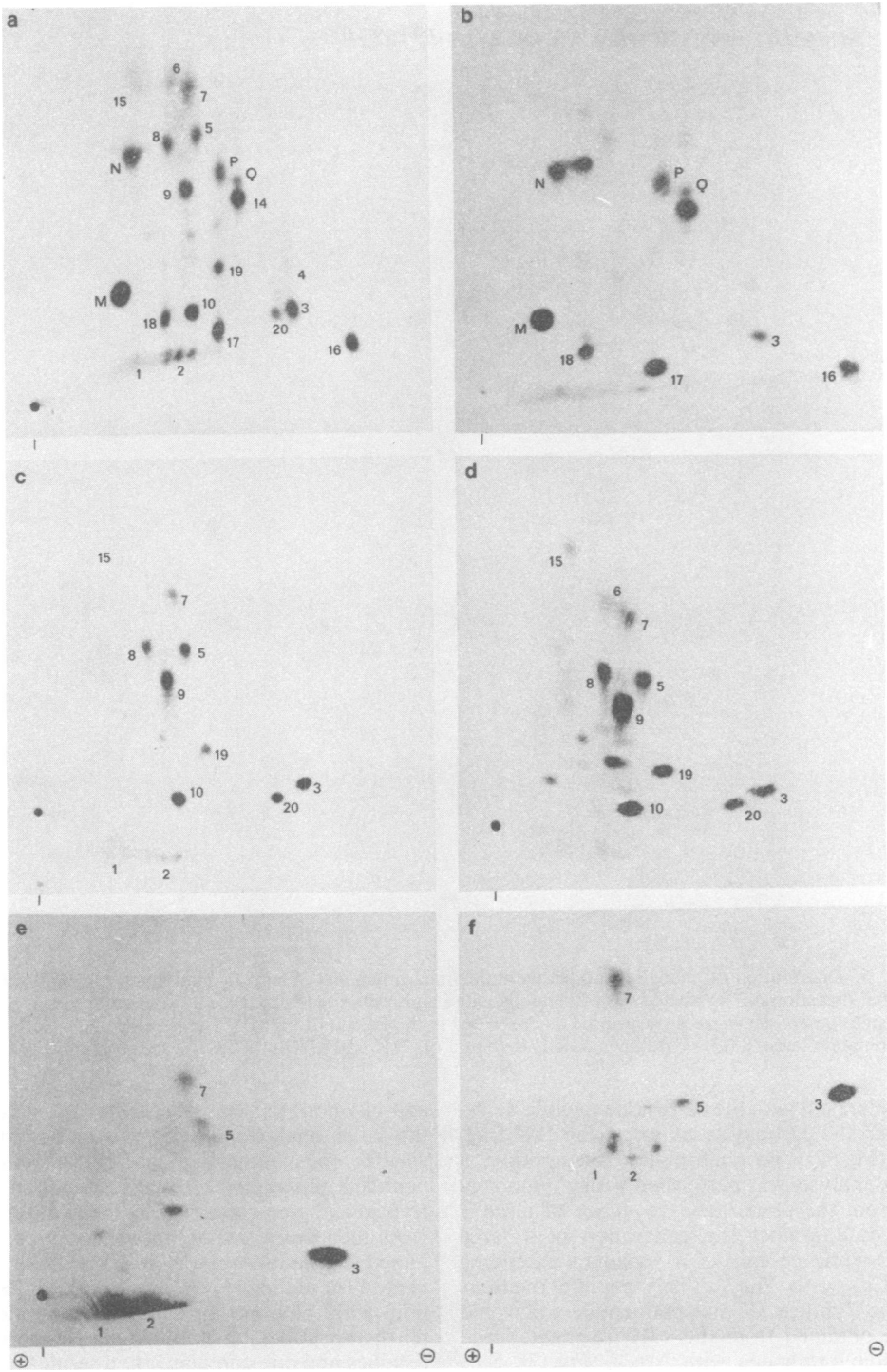


FIG. 4. Fingerprint analysis of fragments of [$^{35}$S]methionine-labeled large T. [$^{35}$S]methionine-labeled large T was extracted from infected cells, immunoprecipitated, and subjected to partial proteolysis as described in the text. The digestion products were separated on preparative polyacrylamide gels, and individual fragments were redigested with trypsin under standard conditions and subjected to peptide mapping on TLC plates. With the exception of panel f, all peptide maps were obtained by electrophoresis in buffer B (1,000 V, 25 min) and chromatography in buffer D (5 h). In the case of panel f, peptide mapping was performed as described in the legend to Fig. 1b, to resolve peptides 1 and 2. Autoradiographs were taken as described previously (3). The nomenclature of the peptides is according to Mann et al. (30) and differs from the numbering of the phosphopeptides in Fig. 1b. Panels: a, complete large T; b, 17K fragment; c, 67K fragment; d, 58K fragment; e and f, 22K fragment.

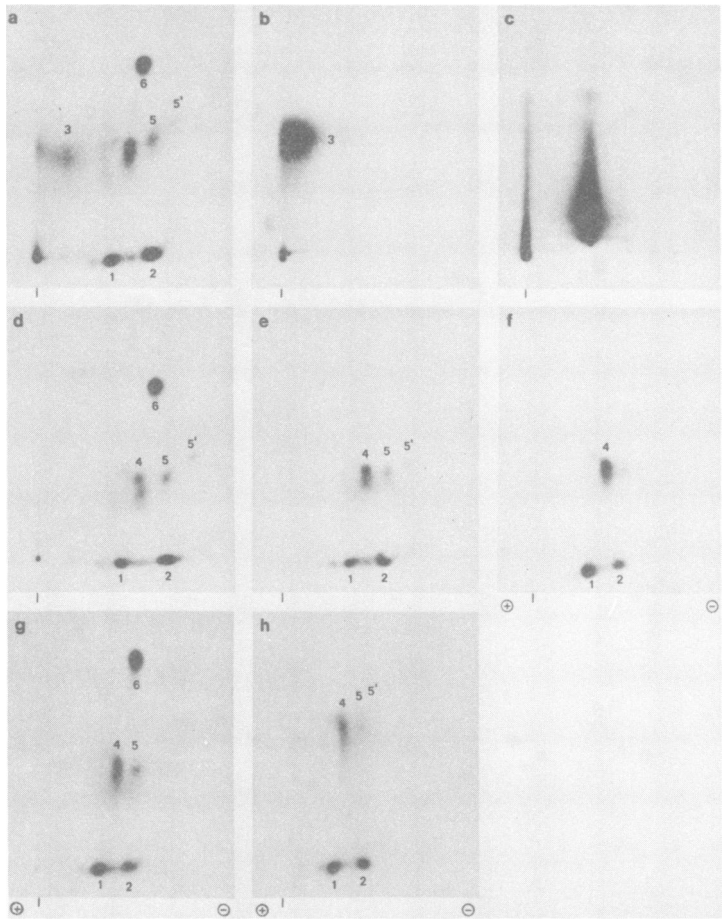


FIG. 5. Distribution of phosphopeptides within defined fragments of large T. Individual fragments obtained by partial digestion of $^{32}P$-labeled large T were isolated and redigested with trypsin (or chymotrypsin, panel c). The resulting peptides were fingerprinted as described in the legend to Fig. 1b. The peptide maps were derived from complete large T (a), or fragments of 17K (b and c), 71K (d), 67K (e), 58K (f), 26K (g), and 22K (h).

3 was identical with the 60-residue peptide deriving from the carboxy terminus of the 17K fragment (Fig. 7). To confirm this assumption, a similar analysis was performed with chymotrypsin. From the amino acid sequence of large T one would predict the generation of a large chymotryptic peptide of 41 residues extending from $Cys_{105}$ to $Phe_{145}$. This peptide overlaps with the tryptic 6.5K phosphopeptide and would not be obtained from the 17K fragment, since the latter terminates with $Arg_{130}$ (Fig. 7). Size analysis of a chymotryptic digest of large T revealed large phosphopeptides of 4.5K to 5.5K and smaller peptides migrating with the front (Fig. 6, lane c). As expected, the 17K fragment did not give rise to a large chymotryptic phosphopeptide (Fig. 6, lane h). These results were consistent with the predictions. Large T, the 17K fragment, and the isolated phosphopeptide 3 were also digested sequentially with trypsin and chymotrypsin or vice versa. In each case, the large phosphopeptides were converted to smaller ones, indicating that the large tryptic peptide(s) contained cleavage sites for chymotrypsin and vice versa (Fig. 6, lanes d, e, and i).

As mentioned above, phosphopeptide 3 contained serine phosphate and threonine phosphate. Theoretically, cleavage of the 6.5K peptide with chymotrypsin could generate two phosphopeptides, one containing two threonine residues and one containing five serines and two threonines. Two-dimensional analysis of chymotryptic digests of the 17K fragment or of the 6.5K peptide revealed that all phosphorylated residues were located in a single peptide (Fig. 5c). This peptide represents a segment that overlaps between the tryptic 6.5K phosphopeptide and the chymotryptic 5.5K phosphopeptide reaching from $Cys_{105}$ to $Lys_{127}$ (Fig. 7).

Phosphopeptide 3′ (Fig. 1b) was also ana-

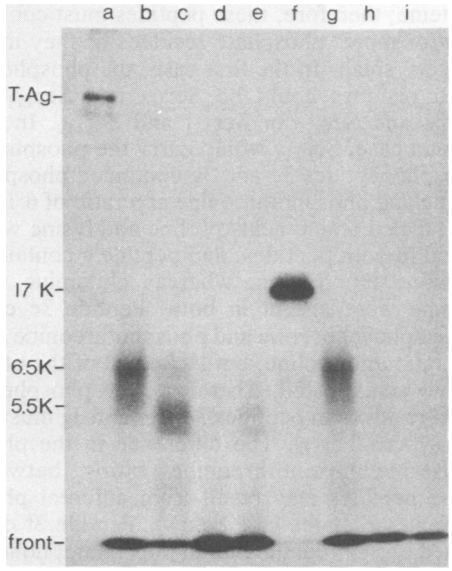


FIG. 6. Analysis of various digests of $^{32}P$-labeled large T by SDS-PAGE. Large T or the 17K fragment were isolated and digested with trypsin or chymotrypsin or sequentially with either enzyme and analyzed on a 10 to 30% polyacrylamide gradient gel. Lanes a through e: large T, untreated (a), digested with trypsin (b) or chymotrypsin (c), or digested sequentially with trypsin and chymotrypsin (d) or chymotrypsin and trypsin (e). Lanes f through i: isolated 17K fragment, undigested (f), digested with trypsin (g) or chymotrypsin (h), or digested with trypsin and chymotrypsin (i).

lyzed. Acid hydrolysis revealed that it contained serine phosphate and threonine phosphate at a 6:1 ratio. On polyacrylamide gels it migrated with the front. During peptide mapping it exhibited properties similar to the chymotryptic peptide generated from the 17K fragment (compare Fig. 1b and 5c). These data suggest that phosphopeptide 3' is a degradation product of phosphopeptide 3.

Since the 17K fragment contained more than half of the phosphate label of large T and the phosphorylated segment from $Cys_{105}$ to $Lys_{127}$ contains five serine and two threonine residues, it was probable that phosphorylation occurred at multiple sites. Consistent with this prediction is the finding of serine phosphate and threonine phosphate at a 6:1 ratio.

Indeed, upon sequential digestion of the 17K fragment with chymotrypsin (or trypsin) and pronase, several new phosphopeptides were generated (Fig. 8), some of which might be overlapping peptides. Two groups of peptides were observed. Peptides of the first group migrated to the anode during electrophoresis at pH 1.9; hence, they must have a net negative charge (Fig. 8, peptides a through d). The other peptides migrated to the cathode and were neutral or positively charged (Fig. 8, peptides e through g). Since carboxyl groups are not dissociated at this pH, the net negative charge of a peptide can result from two or more phosphate residues or from cysteic acid plus phosphate residues. How-

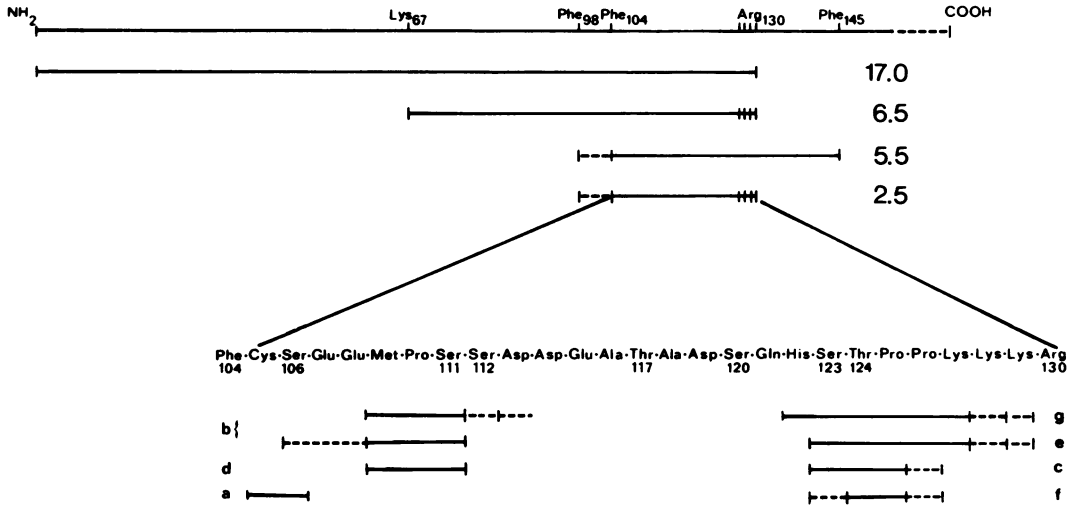


FIG. 7. Schematic representation of the amino-terminal region of large T. The upper line represents part of a linearized molecule of large T. The cleavage sites for trypsin and chymotrypsin that are relevant for the generation of the 17K fragment or the respective phosphopeptides are indicated. The segment overlapping between the tryptic 6.5K and the chymotryptic 5.5K phosphopeptides is shown enlarged with its amino acid sequence (11, 34). The lower part shows the possible locations of the phosphopeptides generated by pronase (Fig. 8) as deduced from their amino acid composition (Table 1).

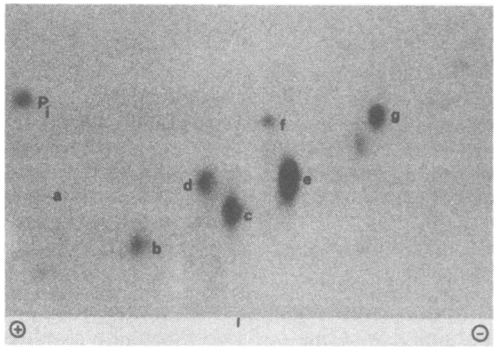


FIG. 8. Fingerprint analysis of pronase peptides generated from the 17K fragment of $^{32}$P-labeled large T. The amino-terminal 17K fragment generated by partial tryptic digestion of $^{32}$P-labeled large T was isolated and redigested with trypsin and pronase. The digestion products were subjected to peptide mapping as described in the legend to Fig. 1b, except that electrophoresis was carried out at 1,500 V for 20 min. The TLC plate was then autoradiographed as described in the text.

ever, phosphoamino acids and di- and tripeptides, though containing only one phosphate group, may also migrate slightly to the anode.

The individual peptides were isolated from TLC plates and analyzed for their contents in phosphoserine and phosphothreonine. Their amino acid composition was determined by radiolabeling large T with amino acids, isolating the 17K fragment and digesting it with pronase, and analyzing the resulting phosphopeptides by peptide mapping. In some cases intact large T was directly digested with trypsin and pronase. The results presented in Table 1 can be summarized as follows. Phosphopeptides 3a, 3b, and 3d contained only serine-bound phosphate. Phosphopeptide 3a contained only cysteine (cysteic acid), but no methionine and proline; therefore, the phosphorylated residue in this peptide must be $Ser_{106}$ (Fig. 7). Peptides 3b and 3d contained methionine and proline, but no alanine and no cysteine; therefore, these peptides must contain two (or more) phosphate residues or they must be very small. In the first case, the phosphorylated residues could be $Ser_{106}$ and $Ser_{111}$ or $Ser_{106}$ and $Ser_{112}$ or $Ser_{111}$ and $Ser_{112}$. In the second case, $Ser_{111}$ would carry the phosphate. Phosphopeptides 3e and 3g contained phosphoserine and phosphothreonine at a ratio of 6:1; of the labeled amino acids proline and lysine were found in both peptides, and peptide g contained histidine in addition, whereas glutamine and alanine were absent in both. Peptide 3c contained phosphoserine and phosphothreonine at a 1:1 ratio and proline, but lacked all of the other amino acids tested. Therefore, the phosphorylated residues in peptides 3c, 3e, and 3g must be $Ser_{123}$ and $Thr_{124}$. The difference in the phosphoserine/phosphothreonine ratios between these peptides may result from different phosphorylation states (see below). Peptide 3f contained only threonine-bound phosphate; none of the other labeled amino acids could be detected. Since it did not comigrate with authentic phosphothreonine it must contain an additional amino acid, possibly unphosphorylated $Ser_{123}$. Presently, we cannot decide whether phosphopeptide 3f contains $Thr_{124}$ or $Thr_{117}$. In summary, the amino-terminal region contains at least four phosphorylation sites which are clustered between $Cys_{105}$ and $Lys_{127}$. $Ser_{106}$, $Ser_{123}$, and $Thr_{124}$ were identified as phosphorylated residues; uncertainty remains about $Ser_{111}$, $Ser_{112}$, and $Thr_{117}$, whereas $Ser_{120}$ is definitely unphosphorylated.

**Identification of phosphopeptides of the carboxy-terminal region.** Four phosphoserine-containing peptides were derived from a fragment extending 26K upstream from the carboxy terminus of large T, presumably up to $Thr_{518}$ (Fig. 5g). To analyze and identify these phosphopeptides, a strategy similar to that described above for the amino-terminal 17K fragment was applied. The size of the phosphopeptides isolated

TABLE 1. Amino acid composition of the amino-terminal phosphopeptides generated by pronase digestion of the 17K fragment

| Peptide | Ala[a] | Cys[b] | Gln[b] | His[a] | Lys[a] | Met[b] | Pro[b] | Tyr[a] | Ser(P)[c,d] | Thr(P)[c] | Ser(P):Thr(P)[a,c] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3a | − | + | − | − | − | − | − | − | + | − | − |
| 3b | − | − | − | − | − | + | + | − | + | − | − |
| 3c | − | − | − | − | − | − | + | − | + | + | 1 |
| 3d | − | − | − | − | − | + | + | − | + | − | − |
| 3e | − | − | − | ± | + | − | + | − | + | + | 6 |
| 3f | − | − | − | − | − | − | − | − | − | + | − |
| 3g | − | − | − | ++ | + | − | + | − | + | + | 6 |

[a] Determined by elution from TLC plates and scintillation spectrometry.
[b] Determined by autoradiography of peptide maps.
[c] Determined by acid hydrolysis of isolated $^{32}$P-labeled peptides.
[d] Phosphorylated forms of amino acids are indicated by (P).

from TLC plates was determined on polyacrylamide gradient gels. Phosphopeptides 1 and 2 migrated with the front, indicating that they are smaller than 3K (the limit of resolution of the gel). Phosphopeptides 4, 5, and 5′ migrated heterogeneously with apparent molecular weights between 4K and 7K. The carboxy-terminal phosphothreonine peptide (which consists of 11 amino acids) migrated with the front (data not shown).

**Characterization of phosphopeptides 1 and 2.** From their behavior during chromatography it can be deduced that peptides 1 and 2 are extremely polar (Fig. 1b). When the electrophoresis on TLC plates was performed at pH 8.9 (where all carboxyl groups are dissociated) peptides 1 and 2 migrated far to the anode (data not shown), indicating that they are very acidic. Determination of the amino acid composition of phosphopeptides 1 and 2 by differential labeling revealed that of 13 different amino acids only alanine, glutamine, and lysine were present in these peptides (Table 2). Based on the amino acid sequence of large T, only one tryptic peptide fulfills these conditions, extending from $Asn_{631}$ to $Lys_{645}$ (Fig. 9) and containing eight acidic residues. The similar physical properties of phosphopeptides 1 and 2 and the labeling data suggested that these peptides contain overlapping sequences. They could differ by charge or by length.

A difference in charge between phosphopeptides 1 and 2 could result from a different degree of phosphorylation; the less positively charged peptide 1 might have both serines phosphorylated, whereas peptide 2 might carry only one phosphate. Alternatively, if both peptides contained the same amount of phosphate, peptide 2 could result from incomplete digestion with trypsin and might contain one basic residue more than peptide 1. To distinguish between differential phosphorylation and incomplete digestion, phosphopeptides 1 and 2 were isolated and cleaved with pronase or acid. If peptides 1 and 2 differed in their phosphate content one would expect different phosphate-containing cleavage products. Upon partial acid hydrolysis (5.6 N HCl, 37°C, 40 h [26]), the same products were generated from both peptides, indicating that they contained the same phosphorylated residues (Fig. 10b and d). Digestion of phosphopeptide 1 with pronase generated one phosphopeptide (designated phosphopeptide 1a in Fig. 10a). Pronase digestion of phosphopeptide 2 revealed predominantly phosphopeptide 2a and a small amount of phosphopeptide 1a (Fig. 10c). The third phosphopeptide visible in Fig. 10c was only obtained when phosphopeptide 2 was first isolated from TLC plates and then treated with pronase, but not when tryptic peptides were pronase digested before peptide mapping. Therefore, we assumed that this peptide and maybe peptide 2a were digestion intermediates. The reason for the limited digestion of the isolated phosphopeptide 2 is not known.

Since peptide 1a migrated to the anode at pH 1.9 it was possible that it contained two phosphate residues. In this case, it had to contain the six acidic residues between $Ser_{632}$ and $Ser_{639}$ (see sequence in Fig. 9). To answer this question phosphopeptide 1a was analyzed by electrophoresis at pH 4.7 and 8.9. As mentioned above, at pH 1.9 only amino and phosphate groups of phosphopeptides are charged; at pH 4.7, all carboxyl groups are dissociated, and at pH 8.9, the phosphate residues acquire a second negative charge. From the difference in migration at pH 1.9, 4.7, and 8.9 (data not shown) we deduced that phosphopeptide 1a does not contain more than one acidic residue and contains only one phosphate group.

To determine whether $Ser_{632}$ or $Ser_{639}$ was phosphorylated, [$^{14}$C]glutamine-labeled large T was digested with trypsin and pronase and analyzed by peptide mapping. Two of the glutamine-containing peptides comigrated with phosphopeptides 1a and 2a. Therefore, the serine next to the glutamine (namely, $Ser_{639}$) must be phosphorylated. Considering that partial acid hydrolysis generated the same products from

TABLE 2. Amino acid composition of the tryptic phosphoserine peptides in the carboxy-terminal region of large T

| Peptide | Ala[a] | Arg[b] | Cys[b] | Gln[b] | His[a] | Ile[a] | Leu[a] | Lys[a] | Met[b] | Pro[b] | Val[a] | Thr[b] | Thr(P)[c] | Ser(P)[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | + | − | − | + | − | − | − | + | − | − | − | − | − | + |
| 2 | + | − | − | + | − | − | − | + | − | − | − | − | − | + |
| 4 | + | − | − | + | − | ±[d] | + | + | ± | − | + | − | − | + |
| 5 | ± | + | + | + | + | + | − | − | + | + | + | + | − | + |
| 5′ | ± | + | + | ++[e] | ++ | ++ | ++ | − | − | ++ | ++ | + | − | + |

[a] Determined by scintillation counting of isolated peptides.
[b] Determined by autoradiography.
[c] Phosphorylated forms of amino acids are indicated by (P).
[d] Slightly above background levels.
[e] Three to six times the amount detected in peptide 5.

$Asn_{515}$ -Lys-Arg▼-Thr-Gln-Ile-Phe-Pro-Pro-Gly-Ile-Val-Thr-Met-Asn-$Glu_{530}$
Tyr-Ser-Val-Pro-Lys▼-Thr-Leu-Gln-Ala-Arg-Phe-Val-Lys-Gln-Ile-Asp-Phe-Arg-Pro-$Lys_{550}$
Asp-Tyr-Leu-Lys-His-Cys-Leu-Glu-Arg▼-Ser-Glu-Phe-Leu-Leu-Glu-Lys▼-Arg▼-Ile-Ile-$Gln_{570}$
Ser-Gly-Ile-Ala-Leu-Leu-Leu-Met-Leu-Ile-Trp-Tyr-Arg-Pro-Val-Ala-Glu-Phe-Ala-$Gln_{590}$
Ser-Ile-Gln-Ser-Arg▼-Ile-Val-Glu-Trp-Lys-Glu-Arg-Leu-Asp-Lys▼-Glu-Phe-Ser-Leu-$Ser_{610}$
Val-Tyr-Gln-Lys▼-Met-Lys▼-Phe▽-Asn-Val-Ala-Met-Gly-Ile-Gly-Val-Leu-Asp-Trp▽-Leu-$Arg_{630}$▼
Asn-Ser-Asp-Asp-Asp-Asp-Glu-Asp-Ser-Gln-Glu-Asn-Ala-Asp-Lys▼-Asn-Glu-Asp-Gly-$Gly_{650}$
Glu-Lys▼-Asn-Met-Glu-Asp-Ser-Gly-His-Glu-Thr-Gly-Ile-Asp-Ser-Gln-Ser-Gln-Gly-$Ser_{670}$
Phe▽-Gln-Ala-Pro-Gln-Ser-Ser-Gln-Ser-Val-His-Asp-His-Asn-Gln-Pro-Tyr▽-His-Ile-$Cys_{690}$
Arg▼-Gly-Phe-Thr-Cys-Phe▽-Lys▼-Lys-Pro-Pro-Thr-Pro-Pro-Pro-Glu-Pro-Glu-Thr-$Thr_{708}$-COOH

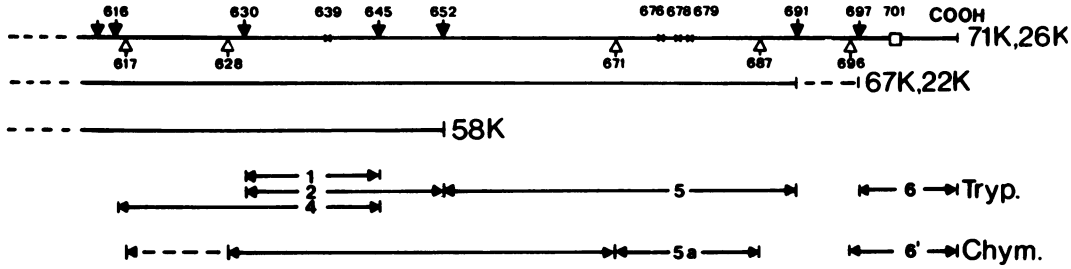


FIG. 9. Representation of the carboxy-terminal region of large T. The upper part represents the predicted amino acid sequence (11, 34) of the 26K fragment. Indicated are those cleavage sites for trypsin (▼) and chymotrypsin (▽) which are referred to in the text. The lower part shows the carboxy-terminal region downstream from $Phe_{617}$ with the extensions of the phosphopeptides that can be generated by digestion with trypsin or with chymotrypsin as deduced from the experimental results in accordance with the amino acid sequence. The phosphorylated $Ser_{639}$ and the possibly phosphorylated $Ser_{676}$, $Ser_{677}$, and $Ser_{679}$ are indicated by ×; the phosphorylated $Thr_{701}$ is marked by □.

both phosphopeptides 1 and 2 and assuming that phosphopeptide 2a was a digestion intermediate (which would give rise to phosphopeptide 1a upon complete digestion), we conclude that both phosphopeptides 1 and 2 are phosphorylated only at $Ser_{639}$. Therefore, their different migration during electrophoresis may reflect incomplete digestion rather than a different degree of phosphorylation. Analysis on a 10 to 37.5% polyacrylamide gradient gel revealed that phosphopeptide 2 was larger than phosphopeptide 1, suggesting that phosphopeptide 2 consisted of phosphopeptide 1 (from $Asn_{631}$ to $Lys_{645}$) plus an extension at one end. The labeling data (Table 2) and the hydrophilic properties argue against an extension on the amino-terminal side (up to $Phe_{617}$), but are compatible with an elongation on the carboxy-terminal side down to $Lys_{652}$ because all additional residues are polar and the second lysine residue would render phosphopeptide 2 more positively charged than phosphopeptide 1. We conclude that phosphopeptides 1 and 2 represent overlapping peptides, peptide 1 extending from $Asn_{631}$ to $Lys_{645}$ and peptide 2 extending from $Asn_{631}$ to $Lys_{652}$, and that both peptides are phosphorylated at $Ser_{639}$.

**Characterization of phosphopeptide 4.** Phosphopeptide 4 appeared on polyacrylamide gels as a relatively large peptide of 4.5 K. It was generated from all carboxy-terminal fragments of large T including the 58K fragment which terminates at $Lys_{652}$. Therefore, we concluded that phosphopeptide 4 must be derived from a region between $Thr_{518}$ (the amino terminus of the 26K fragment) and $Lys_{652}$ (the carboxy terminus of the 58K fragment) (Fig. 9). From the amino acid sequence one can predict that tryptic digestion would generate four serine-containing peptides from this region, if one excludes the acidic phosphopeptide 1. Two of these peptides, extending from $Thr_{518}$ to $Lys_{535}$ and from $Ile_{568}$ to $Arg_{595}$, were identified by differential labeling as methionine peptides 5 and 7, respectively (Fig. 4e and f). The other two peptides, extending from $Ser_{560}$ to $Lys_{566}$ and from $Glu_{606}$ to $Lys_{614}$, respectively, can also be excluded be-

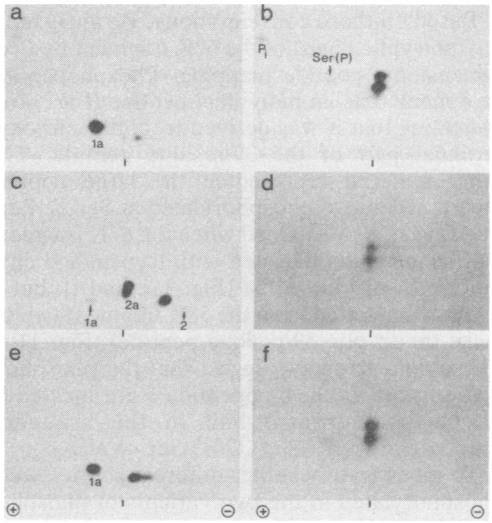


FIG. 10. Analysis of phosphopeptides 1, 2, and 4 by pronase digestion and partial acid hydrolysis. Phosphopeptides 1 (a and b), 2 (c and d), and 4 (e and f) were separated by two-dimensional peptide mapping as described in the legend to Fig. 1b. After elution from the TLC plates, the isolated peptides were subjected to pronase digestion (a, c, and e) or to partial acid hydrolysis (5.6 N HCl, 37°C, 40 h [26]) (b, d, and f). Subsequently, the products were analyzed by peptide mapping as described in the legend to Fig. 8.

cause of their size as well as their amino acid composition. Therefore, we assumed that phosphopeptide 4 represented an incomplete digestion product.

Phosphopeptide 4 was analyzed by cleavage with chymotrypsin or pronase or acid. Digestion with pronase or partial acid hydrolysis yielded the same products (Fig. 10e and f) as obtained from phosphopeptide 1 (Fig. 10a and b), suggesting that peptide 4 contained the same phosphorylation site as peptide 1. Treatment of the isolated peptide 4 with chymotrypsin was rather inefficient, but when large T or carboxy-terminal fragments of large T were sequentially digested with trypsin and chymotrypsin, phosphopeptide 4 disappeared to yield a peptide with properties similar to those of phosphopeptide 2 (phosphopeptide 4a in Fig. 11e and f). These results can be best explained by assuming that phosphopeptide 4 is composed of phosphopeptide 1 plus an extension at the amino terminus of the latter up to $Phe_{617}$ (see sequence in Fig. 9). It does not exceed beyond $Phe_{617}$, because the bonds $Lys_{614} \cdot Met \cdot Lys_{616} \cdot Phe_{617}$ are very efficiently cleaved to yield methionine peptide 3, which represents the dipeptide Met · Lys (25) (Fig. 4e and f). The extension would render phosphopeptide 1 more hydrophobic; upon digestion with chymotrypsin, one would expect cleavage at

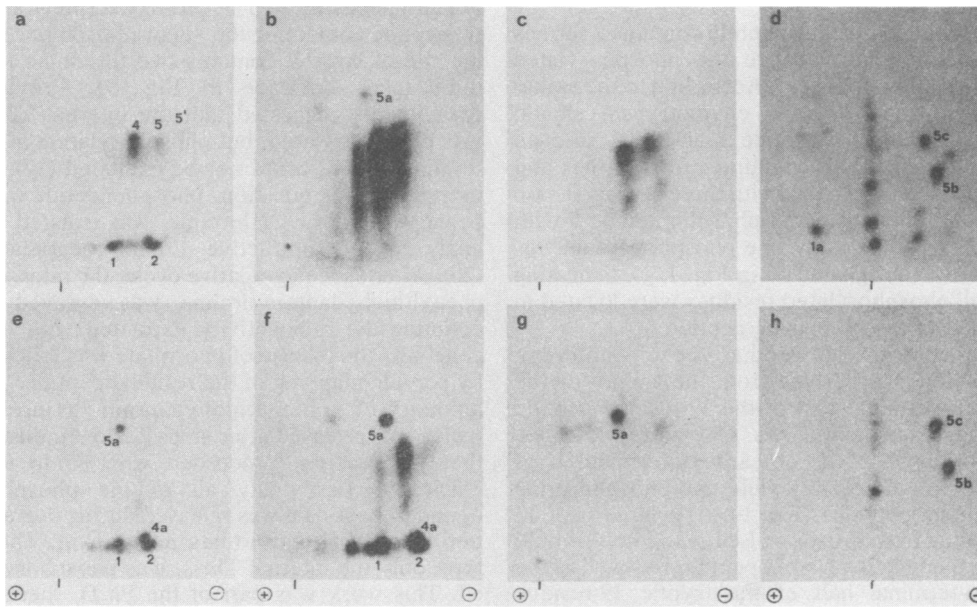


FIG. 11. Analysis of phosphopeptides 4 and 5 by digestion with chymotrypsin or pronase. The 67K and 58K fragments of large T were obtained by partial digestion as described in the text. The isolated fragments were redigested with trypsin or chymotrypsin or sequentially with both enzymes, and the digestion products were analyzed by peptide mapping as described in the legend to Fig. 1b. Panels a, b, e, and f: 67K fragment, digested with trypsin (a), chymotrypsin (b), trypsin and chymotrypsin (e), or chymotrypsin and trypsin (f). Panel c, 58K fragment, digested with chymotrypsin; panel g, tryptic phosphopeptide 5, isolated and redigested with chymotrypsin; panels d and h, phosphopeptides 5 and 5', respectively, isolated and redigested with pronase.

$Trp_{628}$ to yield phosphopeptide 1 extended only by $Arg_{630}$ and $Leu_{629}$. Such a peptide would be slightly more hydrophobic and would contain one positive charge more than phosphopeptide 1. The properties of phosphopeptide 4a (Fig. 11e and f) are consistent with these predictions.

**Characterization of phosphopeptide 5.** The amino acid composition of phosphopeptide 5 as determined in various labeling experiments (Table 2), its apparent molecular weight of approximately 5K (data not shown), and the finding that this peptide was absent on peptide maps of the 58K fragment (Fig. 5f), as was methionine peptide 2 (Fig. 4d), suggested that it was identical with a peptide of 39 amino acids extending from $Asn_{653}$ to $Arg_{691}$. On fingerprints of amino acid-labeled large T this peptide appeared as a characteristic group of three to five peptides (Fig. 4f), two of which comigrated with phosphopeptides 5 and 5'. The various forms of the 39-residue peptide might derive from incomplete digestion or incomplete oxidation of the cysteine residue or from different degrees of phosphorylation. Of the two phosphorylated variants of the 39-residue peptide, phosphopeptide 5 always contained more phosphate label than did peptide 5', whereas peptide 5' contained reproducibly three to six times more amino acid-label than did peptide 5. From this finding we conclude that phosphopeptides 5 and 5' differ in their extent of phosphorylation.

Since the 39-residue peptide contains 7 serines we attempted to localize the phosphorylated residues more precisely. According to the amino acid sequence (Fig. 9) chymotrypsin should cleave the tryptic peptide 5 at $Phe_{671}$ into an amino-terminal half containing four serines and a carboxy-terminal half with three serines. Upon digestion of the isolated phosphopeptide 5 with chymotrypsin, mainly one phosphopeptide was generated (designated 5a in Fig. 11g), indicating that all phosphorylated residues were located in one half of the 39-residue peptide.

The question of whether the chymotryptic peptide 5a was derived from the amino- or the carboxy-terminal part of the 39-residue peptide was addressed as follows. Chymotryptic digestion of large T (or of carboxy-terminal fragments) would possibly yield two phosphoserine containing peptides (Fig. 9): (i) a large peptide extending from $Asn_{618}$ or $Leu_{629}$ to $Phe_{671}$ would cover the tryptic phosphopeptide 1 as well as the amino-terminal half of the tryptic 39-residue peptide; (ii) a second peptide extending from $Gln_{672}$ to $Tyr_{687}$ would represent the carboxy-terminal part of the 39-residue peptide. Thus, if phosphopeptide 5a was derived from the amino-terminal half of the 39-residue peptide, it would be contained in the large chymotryptic peptide, but if it originated from the carboxy-terminal half, it would represent an individual peptide.

Based on these considerations, we analyzed a chymotryptic digest of the 67K fragment by two-dimensional peptide mapping. Phosphopeptide 5a appeared as an individual peptide (Fig. 11b), indicating that it was derived from the carboxy-terminal part of the 39-residue peptide. The other material represented the large peptide mentioned above phosphorylated at $Ser_{639}$. Peptide 5a was also obtained when the 67K fragment was sequentially digested with trypsin and chymotrypsin or vice versa (Fig. 11e and f), but it was not generated from the 58K fragment (which lacks the tryptic 39-residue peptide) (Fig. 11c). These data strongly suggest that the phosphorylation sites in phosphopeptide 5 are located in its carboxy-terminal half in the sequence $Pro_{674}$ · Gln · Ser · Ser · Gln · Ser · $Val_{680}$.

To investigate whether different serines were phosphorylated in the two variants of phosphopeptide 5 (5 and 5'), they were isolated and digested with pronase. Both peptides yielded two major peptides (designated 5b and 5c) and several minor, probably overlapping, peptides (Fig. 11d and h). Phosphopeptide 5 yielded additional peptides, one of which comigrated with phosphopeptide 1a (a digestion product of phosphopeptide 1 [Fig. 10a]). This result further supports the conclusion that phosphopeptides 5 and 5' are differentially phosphorylated.

**Determination of the phosphorylated threonine in phosphopeptide 6.** The carboxy-terminal tryptic peptide containing the major phosphothreonine site of large T contains two threonine residues (see sequence in Fig. 9). Previous experiments suggested that the internal $Thr_{701}$ was phosphorylated, but phosphorylation at the terminal $Thr_{708}$ could not be excluded (37). To investigate this question, phosphopeptide 6, labeled with $^{32}P$ or [$^3$H]proline, was isolated and analyzed by subtractive Edman degradation (20). After each degradative cycle, the release of phenylthiohydantoin-proline was assayed by counting the radioactivity extracted with benzene, and the release of phosphate was followed by peptide mapping of the residual peptide. The amount of phenylthiohydantoin-[$^3$H]proline which was released after steps 2 and 3 indicated that the reaction proceeded with 80 to 90% efficiency. Essentially all of the phosphate bound to peptide 6 was released during degradation steps 4, 5, and 6 (data not shown). Therefore, only the internal $Thr_{701}$ was phosphorylated. This work was part of the Ph.D. thesis of Alexa Kaiser, University of Freiburg.

**Analysis for ADP-ribosylation.** Recently, it has been reported that large T is modified by poly(ADP)-ribosylation (14). We have analyzed the phosphopeptides of large T to determine whether they contained ADP-ribose. SV40-infected TC7 cells were labeled with [$^3$H]adenine (1 mCi per plate). Large T was purified by

immunoprecipitation and SDS-PAGE as described above. After elution from the gel and precipitation with trichloroacetic acid, about 8,000 cpm was associated with large T. However, no adenine label was found in specific association with any of the phosphopeptides described in the present paper. Maybe ADP ribosylation occurs in one of the minor phosphorylation sites which have not yet been analyzed.

## DISCUSSION

We have investigated the number of phosphorylation sites and their location within the polypeptide chain of large T. With a new chromatography system for peptide mapping it was possible to separate six distinct phosphopeptides (including the previously identified carboxy-terminal phosphothreonine peptide [37]). Specific fragments of large T which were generated by partial proteolysis (39) served to determine the origin of the phosphopeptides. To align the fragments within large T, we took advantage of the known arrangement of methionine peptides (25, 30) and the known location of the phosphothreonine peptide (37) of large T.

Fingerprint analysis of individual fragments labeled with [$^{35}$S]methionine or $^{32}$P revealed that a 17K fragment was derived from the amino terminus and that fragments of 71K, 67K, and 58K sharing their amino termini but differing at their carboxy termini were all derived from the unique portion of large T. These results are in agreement with those published by Schwyzer et al. (39). The origin of additional fragments of 26K, 24K, and 22K from the carboxy terminus of large T was deduced from methionine peptide maps and from the presence of the carboxy-terminal phosphothreonine peptide in the 26K fragment. The identification of several peptides by differential labeling allowed us to determine more precisely the termini of the fragments. The 71K fragment extends to the very carboxy terminus of large T, the 67K fragment terminates at $Arg_{691}$ or $Lys_{697}$, and the 58K fragment terminates at $Lys_{652}$. The 26K and 22K fragments have the same carboxy termini as the 71K and 67K fragments, respectively, and extend upstream probably to $Thr_{518}$. It is remarkable that the lack of the carboxy-terminal 56 amino acids in the 58K fragment caused a reduction in apparent molecular weight of 13K compared with the 71K fragment. Obviously, sequences within the 39-residue peptide have a great influence on the migration of large T or fragments thereof in SDS-polyacrylamide gels. Similarly, large T of the deletion mutant 1263 lacks only 15 amino acids within this peptide, but it migrates as if it were smaller by 6K (7, 50).

Mapping of the phosphopeptides revealed that the 17K fragment contained only phosphopeptide 3, whereas all of the other phosphopeptides were derived from a carboxy-terminal fragment of 26K. As reported by others, both the amino-terminal and the carboxy-terminal fragments contained phosphoserine and phosphothreonine (51; M. Kress, personal communication).

**Characteristics of the amino-terminal phosphorylation sites.** The 17K fragment contained more than half of the phosphate label of large T. But this high phosphate content was not reflected on peptide maps where phosphopeptide 3 appeared in quantities similar to those of the other phosphopeptides. The underrepresentation results from its low recovery of 2 to 3% as calculated from radiolabeling with amino acids. The low yields of this peptide made the precise localization of all phosphorylated residues difficult. Analysis of phosphopeptide 3 by SDS-PAGE, sequential digestion with proteases, and two-dimensional peptide mapping revealed that this peptide contained multiple phosphorylation sites, all of which were clustered between $Ser_{106}$ and $Thr_{124}$.

Labeling experiments with various radiolabeled amino acids allowed us to identify $Ser_{106}$, $Ser_{123}$, and $Thr_{124}$ as phosphorylated residues. In addition, at least one of $Ser_{111}$ or $Ser_{112}$ must be phosphorylated as deduced from the amino acid composition of phosphopeptides 3b and 3d, generated by pronase, and their migration to the anode during electrophoresis at pH 1.9. As mentioned above, migration to the anode at this pH indicates either that a peptide has a net negative charge or that it is very small. In the case of phosphopeptides 3b and 3d, a net negative charge can result only from two phosphate residues, because neither peptide contains cysteine (cysteic acid). On the other hand, both peptides must be at least tripeptides, since they contained methionine, proline, and phosphoserine. From their migration during electrophoresis and chromatography it can be deduced that peptide 3b is more negatively charged and more polar than peptide 3d. Furthermore, comparison of the relative migration of phosphoserine, of phosphopeptide 1a [which is a dipeptide composed of Ser (phosphorylated) and Gln; see below], and of phosphopeptide 3d suggests that the last is indeed a tripeptide, composed of Met · Pro · $Ser_{111}$ (phosphorylated), whereas peptide 3b must contain a second phosphate group which may be attached to $Ser_{112}$ or $Ser_{106}$.

The pronase peptides 3c, 3e, and 3g contained both phosphoserine and phosphothreonine, but at different ratios. Since all three peptides contained only $Ser_{123}$ and $Thr_{124}$ as phosphate acceptors, we concluded that these residues might be differentially phosphorylated. The phosphoserine/phosphothreonine ratio of 1:1 in peptide 3c indicated that both residues were fully phosphorylated. The ratio of 6:1 determined for

peptides 3e and 3g can be explained by assuming that each of these peptides represents actually a 6:1 mixture of peptides identical in their overall composition but phosphorylated either at $Ser_{123}$ or $Thr_{124}$. Such peptides would not be separable from each other.

Phosphopeptide 3f contained only threonine-bound phosphate, but from the labeling experiments we could not determine whether this peptide contained $Thr_{117}$ or $Thr_{124}$ as a phosphorylated residue. Considering that $Thr_{124}$ can be phosphorylated without concomitant phosphorylation of $Ser_{123}$, one would assume that peptide 3f contains $Thr_{124}$.

**Properties of the carboxy-terminal phosphorylation sites.** Five tryptic phosphopeptides were derived from the carboxy-terminal 26K fragment. Among these, the phosphothreonine peptide (phosphopeptide 6) has recently been identified as the ultimate tryptic peptide of large T (37, 51). Phosphorylation in this peptide takes place with more than 90% efficiency (37, 38) and is restricted to the internal $Thr_{701}$ as determined by Edman degradation (A. Kaiser, Ph.D. thesis). Phosphopeptides 1, 2, 4, and 5 contained exclusively serine-bound phosphate. Among these, phosphopeptides 1, 2, and 4 represented overlapping peptides containing sequences from $Asn_{631}$ to $Lys_{645}$ (peptide 1) with extensions on the carboxy-terminal side to $Lys_{652}$ (peptide 2) and on the amino-terminal side up to $Phe_{617}$ (peptide 4). Analysis of the isolated peptides by partial acid hydrolysis and by pronase digestion indicated that all of these peptides are phosphorylated at the same serine residue, namely, $Ser_{639}$, located at the end of a cluster of six acidic residues. The assignment of the phosphate to $Ser_{639}$ was based on the generation by pronase digestion of a glutamine-containing peptide (peptide 1a). From the electrophoretic migration at pH 8.9, we deduced that peptide 1a does not contain more than one acidic amino acid. A tripeptide containing $Ser_{639}$ · Gln could have the sequence Asp · Ser · Gln or Ser · Gln · Glu. A peptide comigrating with peptide 1a was also generated by pronase digestion of phosphopeptide 5, which is phosphorylated in the sequence · $Pro_{674}$ · Gln · Ser · Ser · Gln · Ser · Val. Since the pronase peptides generated from phosphopeptides 1 and 5 can correspond only in Ser · Gln (or Gln · Ser), phosphopeptide 1a must be a dipeptide. Digestion of the isolated phosphopeptide 2 with pronase revealed usually three peptides in various ratios; two of these peptides are probably digestion intermediates, because this limited digestion was much less pronounced when a mixture of tryptic peptides was directly digested with pronase (before peptide mapping). The third peptide was identical with phosphopeptide 1a.

It was unexpected that phosphopeptides 1, 2, and 4 contained the same phosphorylation site. On fingerprints, each of these peptides contained amounts of phosphate label similar to those of the phosphothreonine peptide (peptide 6), which also represents one phosphorylation site. Thus, the sum of $^{32}P$ label recovered in peptides 1, 2, and 4 was 2 to 3 times higher than that in peptide 6. Such differences could result from different recoveries of the individual peptides. Since the recovery of the phosphothreonine peptide is relatively high (about 75% [37]), the relative amounts of $^{32}P$ label obtained in $Ser_{639}$ and in $Thr_{701}$ might rather reflect differences in the phosphate metabolism of the respective sites. We have observed that the turnover rates of serine phosphate and threonine phosphate of large T differ by a factor of 3 (38). Moreover, phosphorylation of $Thr_{701}$ in newly synthesized large T requires about 5 h (manuscript in preparation). The slow rates of both threonine phosphorylation and dephosphorylation could result in underrepresentation of the phosphothreonine site, i.e., overrepresentation of phosphoserine sites. Indeed, when $^{32}P$ labeling was performed for 16 h (i.e., 2 half-lives of threonine phosphate) and large T was analyzed by peptide mapping, the most prominent peptide was found to be the phosphothreonine peptide (unpublished data). This observation indicates that phosphate label is incorporated much faster in $Ser_{639}$ than in $Thr_{701}$.

Phosphopeptides 5 and 5′ were identified as two variants of the 39-residue peptide extending from $Asn_{653}$ to $Arg_{691}$, as revealed by size determination, by differential labeling, and by peptide analysis of proteolytic fragments of large T. Phosphopeptides 5 and 5′ appeared on fingerprints only as minor species, partly because of the low recovery of the 39-residue peptide per se and partly because they were phosphorylated to a low extent. From the amino acid label recovered in phosphopeptide 5, in phosphopeptide 5′, and in the unphosphorylated variants we estimate that phosphopeptide 5 may comprise only 5%, and phosphopeptide 5′ may comprise 30 to 50%, of the total amount of the 39-residue peptide. Peptide analysis of chymotryptic digests of large T, of large T fragments, or of the tryptic phosphopeptide 5 suggested that the phosphorylated residue(s) in peptide 5 are located in its carboxy-terminal half. Peptide 5 always contained more phosphate label, but less amino acid label, than peptide 5′. Moreover, pronase digestion of the isolated peptides 5 and 5′ revealed products common to both peptides (peptides 5b and 5c), but additional ones from peptide 5. Both of these findings indicate that peptide 5 is phosphorylated to a higher degree than peptide 5′.

**Differential phosphorylation of large T.** Our

data indicate that at least two segments of large T may exist in different phosphorylation states, one in the amino terminal region involving $Ser_{123}$ and $Thr_{124}$ and the other in the carboxy-terminal region involving at least two of the three serines in the sequence $Ser_{676} \cdot Ser_{677} \cdot Gln \cdot Ser_{679}$. Peptides containing identical phosphate acceptors but differing in their phosphorylation extent must derive from different molecular classes of large T. Several investigators described subspecies of large T differing in their aggregation state (4, 10, 16) or their DNA-binding capacity (16, 32) on one side, or their degree of phosphorylation on the other (15, 31). It will be interesting to investigate how different aggregation forms or DNA-binding species of large T correlate with the phosphorylation state of specific sites.

**Phosphorylation and DNA binding of large T.** The region containing the amino-terminal phosphorylation sites may be involved in the interaction of large T with DNA (42). The phosphorylation sites are located next to a cluster of five basic amino acids which might participate in binding to the DNA backbone. It is conceivable that the introduction of negative charges by phosphorylation modulates the binding affinity of large T. Mann and Hunter described a phosphopeptide which is specifically enriched in chromatin-associated large T (29). (We have not yet analyzed this peptide, because it represented a minor species.) Other investigators reported a positive correlation between the degree of phosphorylation of large T and its affinity for DNA (16, 31); in addition, the DNA-binding capacity of large T was found to be restricted to its oligomeric forms (4), which are phosphorylated to a higher degree than monomers (10, 16). On the other hand, Shaw and Tegtmeyer found no difference in the specific DNA-binding potential between phosphorylated and dephosphorylated large T (41). DNA-binding experiments performed in our laboratory indicate preferential binding of newly synthesized large T which is phosphorylated only to a low extent (manuscript in preparation). However, it is conceivable that DNA binding per se and the regulatory activity of large T in DNA replication and transcription have different requirements with respect to phosphorylation. In vitro transcription systems such as that described by Rio et al. (35) might be useful to clarify this question.

**Phosphorylation and *ts*A mutations.** Walter and Flory (52) reported that phosphorylation is impaired in large T of *ts*A mutants at the restrictive temperature. The basis for this correlation between functional defects and a reduced level of phosphorylation is not yet known. The *ts*A mutants used in those studies map between 0.38 and 0.32 map units on the viral genome (23), whereas the phosphorylation sites are located between 0.520 and 0.508 map units and downstream from 0.218 map units, i.e., outside the mutated region in the *ts*A mutants. This implies that the reduced level of serine phosphorylation does not simply result from a loss of phosphate acceptor sites. An indirect effect of the mutation on phosphorylation could be exerted in several ways. (i) The thermolability of the tertiary structure might favor dephosphorylation. (ii) Some of the serines might be phosphorylated by the protein kinase activity that is associated with large T (1, 4, 17, 48) and that is temperature sensitive in *ts*A mutants (17). (iii) The existence of differentially phosphorylated sites in large T suggests that phosphorylation takes place in a stepwise manner in which each phosphorylation event triggers the subsequent reaction, either by introducing a new microenvironment (the negative charge) or by inducing conformational changes. Either one of these reactions might lead to physically and functionally distinct forms of large T. In the case of *ts* mutants the initial phosphorylation reactions might take place properly, but the thermolability of large T would prevent the subsequent structural changes and the associated phosphorylations.

A recently isolated deletion mutant (*dl*1499) exhibits unusual properties; it is heat sensitive for lytic growth but cold sensitive for transformation (33). Large T of *dl*1499 lacks sequences including $Ser_{657}$ and the phosphorylated $Ser_{639}$. This mutant would provide a useful tool to study the role of this phosphorylation site.

A possible role of phosphorylation in transport and accumulation of large T in the nucleus has been suggested from the properties of a transport-defective large T mutant having a reduced level of phosphorylation (24). It is possible now to investigate which sites are phosphorylated in this mutant.

**General features of the phosphorylation sites of large T.** Phosphorylation sites in proteins are often found in β-turn structures (43). Four of the identified sites in large T may fulfill this rule, i.e., they have a high probability of assuming β-turn conformation (6): $\cdot Pro \cdot Ser_{111} \cdot Ser_{112} \cdot Asp \cdot$, $\cdot Ser_{123} \cdot Thr_{124} \cdot Pro \cdot Pro \cdot$ and $\cdot Pro \cdot Pro \cdot Thr_{701} \cdot Pro \cdot$. Several of the phosphorylated residues are located within acidic environments: $\cdot Ser_{106} \cdot Glu \cdot Glu \cdot$, $\cdot Ser_{111} \cdot Ser_{112} \cdot Asp \cdot Asp \cdot Glu \cdot$, and $\cdot Glu \cdot Asp \cdot Ser_{639} \cdot Gln \cdot Glu \cdot$. Similar phosphorylation site sequences have been found in casein (5) and in an acidic nucleolar protein (28). Sequences of the type $Ser \cdot X \cdot Ac$, where X may be any amino acid and Ac is an acidic amino acid, are recognized by so-called casein kinases (5) or acidic nuclear kinases (2, 18). On the other hand, the carboxy-terminal threonine site resembles sequences recognized by the cyclic nucleotide-

dependent kinases (21). Interesting in this respect is the occurrence of differential phosphorylation in the sequence · Gln · $Ser_{676}$ · $Ser_{677}$ · Gln · $Ser_{679}$ · Val. It is tempting to speculate that $Ser_{679}$ is the first residue to be phosphorylated to yield phosphopeptide 5'. Introduction of the acidic residue might then provide a new recognition site for an acidic kinase leading to subsequent phosphorylation of $Ser_{677}$ or $Ser_{676}$ (or both) to yield phosphopeptide 5. Remarkable is the frequency with which the sequence · Ser · Gln · or · Gln · Ser · occurs in the carboxy-terminal region, but the significance of this amino acid arrangement with respect to phosphorylation is not known.

The amino-terminal phosphorylation sites lie within, but the carboxy-terminal sites are outside the homologous region between SV40 and polyoma virus large Ts. Polyoma virus large T is also serine rich in a region corresponding to the phosphorylation region in SV40 large T (12, 44). It will be interesting to see whether polyoma large T is phosphorylated in the same region or at homologous sites.

## ACKNOWLEDGMENTS

We thank R. Himmelsbach for typing this manuscript and M. Kress for providing his manuscript before publication.

This investigation was supported by the Deutsche Forschungsgemeinschaft.

## LITERATURE CITED

1. **Baumann, E. A., and R. Hand.** 1979. Protein kinase activity associated with the D2 hybrid protein related to simian virus 40 T antigen: some characteristics of the reaction products. Proc. Natl. Acad. Sci. U.S.A. **76**:3688–3692.
2. **Baydoun, H., J. Hoppe, and K. G. Wagner.** 1981. Quaternery structure and ligand specificity of two messenger-independent protein kinases from porcine liver nuclei, p. 1095–1108. *In* O. M. Rosen and E. G. Krebs (ed.), Cold Spring Harbor conferences on cell proliferation, vol. 8: protein phosphorylation. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
3. **Bonner, W. M., and J. D. Stedman.** 1978. Efficient fluorography of $^{3}H$ and $^{14}C$ on thin layers. Anal. Biochem. **89**:247–256.
4. **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.
5. **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.
6. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
7. **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.
8. **Eckhart, W., M. A. Hutchinson, and T. Hunter.** 1979. An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. Cell **18**:925–933.
9. **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 *ts*A mutants of simian virus 40. J. Virol. **29**:753–762.
10. **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.
11. **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.
12. **Friedmann, T., A. Esty, P. La Porte, and P. Deininger.** 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell **17**:715–724.
13. **Gibson, W.** 1974. Polyoma virus proteins: a description of the structural proteins of the virion based on polyacrylamide gel electrophoresis and peptide analysis. Virology **62**:319–336.
14. **Goldman, N., M. Brown, and G. Khoury.** 1981. Modification of SV40 T antigen by poly-ADP-ribosylation. Cell **24**:567–572.
15. **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.
16. **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.
17. **Griffin, J. D., G. Spangler, and D. M. Livingston.** 1979. Protein kinase activity associated with simian virus 40 T antigen. Proc. Natl. Acad. Sci. U.S.A. **76**:2610–2614.
18. **Hoppe, J., and H. Baydoun.** 1981. Substrate specificity of the nuclear protein kinase NII from porcine liver: studies with casein variants. Eur. J. Biochem. **117**:585–589.
19. **Jessel, D., T. Landau, J. Hudson, T. Lalor, D. Tenen, and D. M. Livingston.** 1976. Identification of regions of the SV40 genome which contain preferred T antigen binding sites. Cell **8**:535–545.
20. **Konigsberg, W.** 1967. Subtractive Edman degradation. Methods Enzymol. **11**:461–469.
21. **Krebs, E. G., and J. A. Beavo.** 1979. Phosphorylation-dephosphorylation of enzymes. Annu. Rev. Biochem. **48**:923–959.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
23. **Lai, C.-J., and D. Nathans.** 1975. A map of temperature-sensitive mutants of simian virus 40. Virology **66**:70–81.
24. **Lanford, R. E., and J. S. Butel.** 1980. Biochemical characterization of nuclear and cytoplasmic forms of SV40 tumor antigens encoded by parental and transport-defective mutant SV40-adenovirus 7 hybrid viruses. Virology **105**:314–327.
25. **Linke, H. K., T. Hunter, and G. Walter.** 1979. 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. J. Virol. **29**:390–394.
26. **Light, A.** 1967. Partial acid hydrolysis. Methods Enzymol. **11**:417–421.
27. **Maizel, J. V., Jr.** 1971. Polyacrylamide gel electrophoresis of viral proteins, p. 179–246. *In* K. Maramorosh and H. Koprowski (ed.), Methods in virology, vol. 5. Academic Press, Inc., New York.
28. **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.
29. **Mann, K., and T. Hunter.** 1981. Phosphorylation of SV40 large T antigen in SV40 nucleoprotein complexes. Virology **107**:526–532.
30. **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.
31. **Montenarh, M., and R. Henning.** 1980. Simian virus 40 T-antigen phosphorylation is variable. FEBS Lett. 114:107–110.
32. **Oren, M., E. Winocour, and C. Prives.** 1980. Differential affinities of simian virus 40 large tumor antigen for DNA. Proc. Natl. Acad. Sci. U.S.A. 77:220–224.
33. **Pintel, D., N. Bouck, G. DiMayorca, B. Thimmappaya, B. Swerdlow, and T. Shenk.** 1980. SV40 mutant tsA 1499 is heat-sensitive for lytic growth but generates cold-sensitive rat-cell transformants. Cold Spring Harbor Symp. Quant. Biol. 39:305–309.
34. **Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, C. L. Celma, and S. M. Weissmann.** 1978. The genome of simian virus 40. Science 200:494–502.
35. **Rio, D., A. Robbins, R. Myers, and R. Tjian.** 1980. Regulation of simian virus 40 early transcription *in vitro* by a purified tumor antigen. Proc. Natl. Acad. Sci. U.S.A. 77:5706–5710.
36. **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.
37. **Scheidtmann, K. H., A. Kaiser, A. Carbone, and G. Walter.** 1981. Phosphorylation of threonine in the proline-rich carboxy-terminal region of simian virus 40 large T antigen. J. Virol. 38:59–69.
38. **Scheidtmann, K. H., A. Kaiser, and G. Walter.** 1981. Phosphorylation sites in SV40 large T antigen, p. 1273–1279. *In* O. M. Rosen and E. G. Krebs (ed.), Cold Spring Harbor conferences on cell proliferation, vol. 8: protein phosphorylation. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
39. **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. 225:5627–5634.
40. **Shalloway, D., T. Kleinberger, and D. M. Livingston.** 1980. Mapping of SV40 DNA replication origin region binding sites for the SV40 T antigen by protection against exonuclease III digestion. Cell 20:411–422.
41. **Shaw, S. B., and P. Tegtmeyer.** 1981. Binding of dephosphorylated A protein to SV40 DNA. Virology 115:88–96.
42. **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.
43. **Small, D., P. Y. Chou, and G. D. Fasman.** 1977. Occurance of phosphorylated residues in predicted β-turns: implications for β-turn participation in control mechanisms. Biochem. Biophys. Res. Comm. 79:341–346.
44. **Soeda, E., J. R. Arrand, N. Smolar, and B. Griffin.** 1979. Sequence from the early region of polyoma virus DNA containing the viral replication origin and encoding small, middle and (part of) large T antigens. Cell 17:357–370.
45. **Tegtmeyer, P., B. Andersen, S. B. Shaw, and G. van Wilson.** 1981. Alternative interactions of the SV40 A protein with DNA. Virology 115:75–87.
46. **Tegtmeyer, P., K. Rundell, and J. K. Collins.** 1977. Modification of simian virus 40 protein A. J. Virol. 21:647–657.
47. **Tjian, R.** 1978. The binding site on SV40 DNA for a T-antigen-related protein. Cell 13:165–179.
48. **Tjian, R., and A. Robbins.** 1979. The SV40 tumor antigen is an enzyme. Cold Spring Harbor Symp. Quant. Biol. 44:103–111.
49. **Tooze, J.** 1980. Lytic cycle of SV40 and polyoma virus, p. 125–204. Transformation by SV40 and polyoma virus, p. 205–296. *In* J. Tooze (ed.), Molecular biology of tumor viruses, part 2; DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
50. **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.
51. **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.
52. **Walter, G., and P. J. Flory, Jr.** 1979. Phosphorylation of SV40 large T antigen. Cold Spring Harbor Symp. Quant. Biol. 44:165–169.