MOLECULAR AND CELLULAR BIOLOGY, Dec. 1989, p. 5525–5536
0270-7306/89/125525-12$02.00/0

Vol. 9, No. 12

Properties of the DNA-Binding Domain of the Simian Virus 40 Large T Antigen

DUNCAN McVEY,* MICHAEL STRAUSS,† AND YAKOV GLUZMAN‡
Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

Received 5 May 1989/Accepted 3 September 1989

T antigen (Tag) from simian virus 40 binds specifically to two distinct sites in the viral origin of replication and to single-stranded DNA. Analysis of the protein domain responsible for these activities revealed the following. (i) The C-terminal boundary of the origin-specific and single-strand-specific DNA-binding domain is at or near amino acid 246; furthermore, the maximum of these DNA-binding activities coincides with a narrow C-terminal boundary, spanning 4 amino acids (246 to 249) and declines sharply in proteins with C termini which differ by a few (4 to 10) amino acids; (ii) a polypeptide spanning residues 132 to 246 of Tag is an independent domain responsible for origin-specific DNA binding and presumably for single-stranded DNA binding; and (iii) a comparison of identical N-terminal fragments of Tag purified from mammalian and bacterial cells revealed differential specificity and levels of activity between the two sources of protein. A role for posttranslational modification (phosphorylation) in controlling the DNA-binding activity of Tag is discussed.

Simian virus 40 (SV40) large T antigen (Tag) is a multifunctional DNA-binding protein necessary for the productive infection of permissive cells and for the malignant transformation of nonpermissive cells (72). During the lytic life cycle of SV40, Tag is essential for the autoregulation of early transcription (26, 50, 71) and replication of the viral genome (59, 70). Both of these activities require Tag to interact directly with sequences found within the SV40 early promoter and origin of replication. This region contains two unique DNA-binding sites that differ in the organization of the pentanucleotide sequence 5'-GAGGC (16, 30), which is recognized by Tag. Site I encompasses a 17-base-pair high-affinity binding site composed of two pentanucleotides separated by a 7-base-pair adenine-plus-thymidine-rich spacer (53), and conformation of this site is temperature sensitive (52). Binding of Tag to this site is required primarily for autoregulation of early transcription (51) but also provides an auxiliary function for viral replication (38, 67). Site II consists of four pentanucleotides which form a perfect 27-base-pair palindrome and is an essential component of the core origin of replication (13, 14, 16, 22, 38, 67), which is also involved in the autoregulation of early transcription (51).

The ability of Tag to bind to these sites is defined by a small region of the linear sequence of the protein whose activity is modulated by posttranslational modifications and cofactors. A variety of genetic and biochemical studies have identified the region of Tag responsible for interacting with the viral origin of replication. Mutations clustered between residues 140 and 220 severely reduced the ability of Tag to bind to the origin of replication (33, 48, 49). In addition, when virus harboring a defective origin was used to select second-site revertants, these revertants mapped to amino acids 157 and 166 (40, 59). Biochemical analysis of Tag by partial proteolysis of the protein in solution or when bound to DNA yielded a protein domain spanning residues 131 to 371 which retained the ability to bind SV40 origin sequences (61). Extended exposure of the Tag-DNA complex to high concentrations of protease resulted in a peptide spanning amino acids 140 to 280 remaining associated with the origin (62). The carboxy-terminal boundary of the origin-binding domain has been further refined to amino acid 259 (1, 68), suggesting that the DNA-binding domain of Tag lies between amino acids 130 and 259. Distal to residue 259 is a potential metal-binding motif spanning amino acids 302 to 320 that is required for DNA-binding activity in a variety of proteins (4), but this motif is dispensable for origin-specific DNA-binding activity by Tag.

The specific DNA-binding activity of large Tag is modulated by a variety of effectors. One such effector is posttranslational phosphorylation of Tag. The phosphorylated residues of Tag reside in two clusters, one each at the amino and carboxy termini of the protein (55), and both lie outside the DNA-binding domain. Phosphorylation of these residues affects the ability of Tag to bind origin sequences (34, 44, 63) and support replication of SV40 origin-containing DNA (24, 44). Partial dephosphorylation of Tag with calf intestinal alkaline phosphatase stimulated replication activity in vitro (24, 44) and was paralleled by an increased binding to site II (35, 44, 63), an essential element in the SV40 origin of replication. Furthermore, monomers of Tag are active in DNA binding (9, 19, 20) and contain a different phosphorylation pattern than oligomeric forms of the protein (56). Genetic studies corroborate these biochemical analyses. Point mutations introduced at several of the phosphorylated residues alter the ability of Tag to bind specific DNA sequences in the SV40 origin of replication and to replicate SV40 origin-containing DNA in vivo (57). Like phosphorylation, ATP modulates the DNA-binding activity of Tag (5, 15). The nucleotide-binding domain of the protein is located in the carboxy-terminal portion of the molecule (8), far removed from the region of the protein responsible for physically interacting with the SV40 origin of replication. ATP induces the assembly of a multimeric nucleoprotein complex over the core origin of replication (5, 12, 15).

* Corresponding author.
† Present address: Zentralinstitut für Molekularbiologie, Akademie der Wiessenshaft der DDR, 115 Berlin-Buch, German Democratic Republic.
‡ Present address: Lederle Laboratories, American Cyanamid Co., Pearl River, NY 10965.

Chemical and enzymatic protection studies have revealed that assembly of this structure is accompanied by limited unwinding of the early imperfect palindrome and structural deformation of the A+T-rich region (6), both of which reside in the core origin of replication (14, 38, 67).

In addition to recognizing origin-specific double-stranded DNA sequences, Tag is involved in limited unwinding of the viral origin, which generates regions of single-stranded DNA (74). The ability of Tag to function as a helicase (65) requires it to interact efficiently with single-stranded DNA. However, relatively little is known about its single-stranded DNA-binding activity. Early studies were performed with a replication-defective mutant of Tag (SV80) which possesses reduced origin-specific DNA-binding activity (19, 21, 64). Experiments with nondefective protein demonstrated that single-stranded DNA is an inefficient competitor of origin-specific DNA-binding activity (3). These experiments, however, do not exclude the possibility that both origin-specific and single-stranded DNA-binding activity reside in the same protein domain.

The interaction of Tag with DNA is complex and responds to a variety of factors. The biological consequences of these interactions are diverse but tightly integrated. To understand better the structural and functional interactions of these activities, we defined in detail the intrinsic properties of this DNA-binding domain.

MATERIALS AND METHODS

Construction of *Bal* 31 deletion series. The starting material used in the construction of the deletion series was a clone capable of expressing the amino-terminal 259 amino acids of Tag in *Escherichia coli*. The clone was described in detail previously (68). Briefly, a cDNA of Tag whose ATG had been converted to a *Nde*I site by site-directed mutagenesis was directionally cloned (*Nde*I-*Bam*HI into a modified pAR3038 plasmid [69]) 3′ of the bacteriophage T7 gene 10 promoter and 5′ of the T7 transcription termination signal (43b). All Tag sequences distal to nucleotide 4040 were deleted, and an *Xba*I stop linker (CTAGTCTAGACTAG) was introduced. A unique *Bam*HI site is adjacent and 3′ to the stop linker. The resulting clone served as the starting material for our set of *Bal* 31 deletions. This clone was restricted with *Bam*HI, treated with *Bal* 31, and phenol extracted to stop the exonuclease reaction. The ends of the DNA were filled with Klenow fragment of DNA polymerase I and all four deoxynucleotides. End-filled DNA was incubated with nonphosphorylated *Xba*I stop linker and T4 ligase. Competent *E. coli* DH5 was transformed and screened for plasmid containing the *Xba*I linker. A number of clones were sequenced from the *Xba*1 stop linker by the chemical degradation technique (41). Transcriptional terminators were reconstituted into selected clones, thereby reintroducing the unique *Bam*HI site from the vector next to the *Xba*I stop linker. These clones are isogenic with respect to vector sequences found in the starting material.

Construction of recombinant adenoviruses. The adenovirus vectors used contained the coding region of the truncated Tags with the Tag intron reintroduced, starting from the *Hin*dIII site (nucleotide 5171), covering the *Xba* stop linker, and ending with a *Bam*HI site derived from pT7Tag. This *Bam*HI site was fused to the *Bcl*I site of the SV40 *Bcl*I-*Bam*HI fragment, which contains the polyadenylation signal. The *Hin*dIII-*Bam*HI fragment was expressed behind the major late promoter and one late mRNA leader of adenovirus type 5 after insertion into early region I of a nondefective adenovirus vector (Y. Gluzman, manuscript in preparation).

Protein production. Tag from a mammalian source was produced in HeLa cell suspension cultures (5×10^5/ml) infected with a mixture of wild-type adenovirus type 5 (as a helper) and a defective adenovirus carrying the coding sequence of SV40 Tag in early region I at multiplicities of infection of 10 and 100, respectively. Cells were harvested 42 to 44 h postinfection, and extracts were prepared as described previously (60). Approximately 100 μg of truncated protein can be purified from 4 liters of HeLa cells.

E. coli-produced Tag was prepared in the following manner. Cultures of BL21(DE3) Lys S carrying the expression plasmid were grown at 37°C in LB medium containing 50 μg of ampicillin per ml. When the optical density at 60 nm reached 0.5, isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 0.4 mM and the flasks were kept in a shaking incubator at 37°C for an additional 2 h. Bacteria (500 ml) were harvested by centrifugation and suspended in buffer C (5 ml) (50 mM Tris hydrochloride [pH 7.4], 500 mM LiCl, 1 mM EDTA, 10% glycerol). Phenylmethylsulfonyl fluoride was added to a final concentration of 1 mM, and the suspension was lysed under 1,500-lb/in^2 nitrogen in a French press. After centrifugation at 76,800 × *g* (r_{avg}), the clarified extract was loaded onto a Sephadex G-50 column (3.5 by 47.5 cm) equilibrated in buffer C and developed at a flow rate of 12 ml/h. This step allowed some separation of the full-length protein from shorter contaminants. Fractions containing full-length protein were loaded sequentially onto a protein A-Sepharose column followed by a protein A-Sepharose column onto which the monoclonal antibody PAb419 (27) had been coupled (60). The column was then washed and eluted as described previously (60). Protein concentrations were determined by commercially available Bradford colorimetric dye (Bio-Rad Laboratories, Richmond, Calif.) following the instructions of the manufacturer. Bovine serum albumin was the protein standard used. All the amino-terminal truncated proteins of 246 amino acids and larger yielded 1 to 2 mg/liter of cells.

DNA binding reactions. Plasmid DNA templates containing site I (pOS-1; SV40 nucleotides 5171 to 5228), site II (pSVOd13; SV40 nucleotides 5209 to 128), and the wild-type origin (pSVO+; SV40 nucleotides 5171 to 128) have been described previously (67, 68). An equimolar mixture of these plasmids was assembled and digested to completion with *Taq*I. This enzyme releases intact SV40 origin sequences and generates several plasmid-derived fragments. The mixture was labeled with the Klenow fragment of DNA polymerase I, extracted with phenol and chloroform, and precipitated with ethanol. This mixture (50 ng) was incubated with various amounts of each Tag in origin binding buffer (10 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid] [pH 7.4], 100 mM KCl, 1 mM $MgCl_2$, 5% glycerol, 50 μg of bovine serum albumin per ml) with a final volume of 50 μl for 1 h at 0°C. Purified PAb419 (5 μg) (27) was added, and the incubation continued for an additional 60 min. A 10% (vol/vol) protein A-Sepharose slurry (100 μl) in NET buffer (50 mM Tris hydrochloride [pH 7.4], 150 mM NaCl, 5 mM EDTA, 0.05% Nonidet P-40) was then added, and the reaction mixtures were incubated on a rocker for 50 min at 4°C. The beads were pelleted, washed three times with 1 ml of NET buffer, suspended in 1% sodium dodecyl sulfate (SDS)–25 mM EDTA, incubated at 65°C for 15 min, and electrophoresed on a 6% native polyacrylamide gel in a Tris-borate-EDTA buffer system. Gels were dried onto DE81 paper (Whatman, Inc., Clifton, N.J.) and exposed to Kodak XAR film.

To assay DNA-binding activity directly from bacterial

lysates, we used the following procedure. Growth conditions, times of induction, and harvest of the *E. coli* producing Tag were essentially as described above. Isopropyl-β-D-thiogalactopyranoside-induced cells (10 ml) were concentrated by centrifugation, and the pellet was washed with 4°C 50 mM Tris hydrochloride [pH 8]–10 mM EDTA, transferred to a microcentrifuge tube, and repelleted. The pellet was suspended in 333 μl of 10 mM Tris hydrochloride [pH 8]–1 mM EDTA–50 mM NaCl–0.05% Nonidet P-40 and sonicated (three times for 30 s each on ice). Lysates were clarified by centrifugation. Tag was immunoprecipitated with 9 μg of antibody (1 h on ice) and 100 μl of 10% protein A-Sepharose beads (rocking at 4°C for 50 min). The beads were washed once with NET buffer, followed by a wash with 50 μl of origin binding buffer. The beads were incubated with 50 ng of DNA containing an equimolar mixture of site I, site II, and wild-type origin fragments, as described above, in 20 μl of origin binding buffer. The pellet was gently agitated and allowed to incubate for 1 h on ice. The beads were pelleted, washed three times with NET buffer, suspended in 1% SDS–25 mM EDTA, incubated at 65°C for 15 min, and electrophoresed on a 6% native polyacrylamide gel in a Tris-borate-EDTA buffer system. Gels were dried onto DE81 paper and exposed to Kodak XAR film.

The nitrocellulose filter binding assay used to generate the gel shown in Fig. 4A employed conditions identical to those used above with the purified proteins. The reaction mixtures that did not receive PAb419 were incubated for the same period as the reaction mixtures that did. Tag-DNA complexes were collected by filtration onto nitrocellulose and washed three times with NET buffer. To elute the DNA, we treated each filter with 10 mM Tris hydrochloride [pH 7.4]–5 mM EDTA–0.2% SDS–1 mg of predigested pronase per ml for 2 h at 37°C. Carrier tRNA was added before the eluate was extracted with phenol and chloroform. After being ethanol precipitated, dried, and suspended in 10 mM Tris (pH 7.4)–1 mM EDTA, the DNA was electrophoresed on a 6% native acrylamide gel. The alkaline-treated nitrocellulose used in these experiments was prepared as described in reference 42. Before use, the filter was washed three times with origin binding buffer lacking bovine serum albumin.

To determine relative affinities for the wild-type origin fragment by nitrocellulose filter binding, we incubated various amounts of purified Tag with 1 ng of gel-purified end-labeled wild-type fragment in 30 μl of origin binding buffer. The wild-type fragment was derived from a *Taq*I digest of pSVO+ as described above. The reaction mixtures were incubated for 1 h on ice and slowly filtered through alkaline-treated nitrocellulose. After drying, the filters were counted in liquid scintillant.

Heteroduplex DNA binding assay. mpSV2 contains the *Hin*dIII-*Hpa*II SV40 fragment (nucleotides 5171 to 346) cloned into *Sma-Hin*dIII-digested M13mp9. A synthetic 42-base oligonucleotide (5′-GTTTTCCCAGTCACGACGT TGTAAAACGACGGCCAGTGAATT-3′) was synthesized on a DNA synthesizer (model 380-A; Applied Biosystems) and purified by electrophoresis in a denaturing polyacrylamide gel. After adsorption onto a reverse-phase C18 Sep-Pak cartridge (Waters Associates, Inc., Milford, Mass.), the oligonucleotide was eluted with 60% methanol, lyophilized to dryness, and suspended in H_2O. Purified oligonucleotide (50 ng) was annealed to 1.5 μg of mpSV2 in 10 mM HEPES (pH 7.4)–50 mM NaCl at 50°C for 1 h. After slow cooling to room temperature, the annealed template was labeled with the Klenow fragment of DNA polymerase I (in the presence of [α-^{32}P]dCTP) and [α-^{32}P]dGTP and sequentially extracted with both phenol (twice) and chloroform. The labeled DNA was loaded onto a column constructed from a 1-ml syringe containing Sepharose CL-4B. The column was equilibrated and developed in 10 mM HEPES (pH 7.4)–50 mM NaCl, and peak fractions containing labeled DNA were collected and ethanol precipitated, dried, and resuspended in 10 mM HEPES (pH 7.4)–50 mM NaCl so that the final concentration was 10 ng/μl. Labeled heteroduplex DNA (20 ng) was incubated with various amounts of Tag in 30 μl of origin binding buffer for 1 h on ice. The protein-DNA complexes were collected by filtration onto alkaline-treated nitrocellulose, washed three times with NET buffer, dried, and assayed for radioactivity.

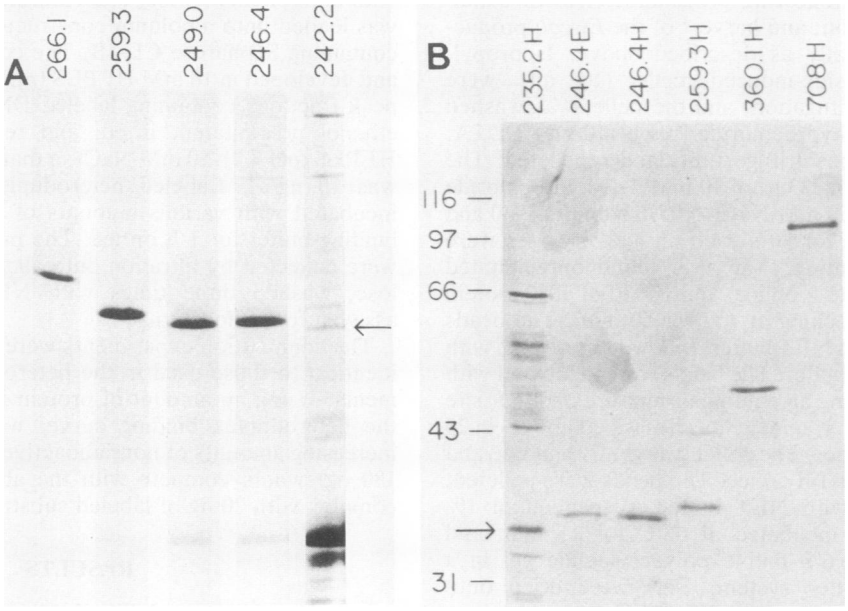
The competition experiments were done under conditions identical to those used in the heteroduplex binding experiments. Using an amount of protein in the linear portion of the heteroduplex binding curve, we determined whether increasing amounts of nonradioactive mpSV2 (0, 20, 60, and 180 ng) would compete with the ability of the protein to complex with 20 ng of labeled substrate.

RESULTS

Isolation of truncated Tags from procaryotic and mammalian cells. The amino-terminal 259 amino acids of SV40 large Tag were previously shown to bind to the SV40 origin of DNA replication in a sequence-specific manner (68). The specific DNA-binding activity of this 259-amino-acid protein was found to be greater than those of larger amino-terminal proteins containing 266 and 272 amino acids (68). To determine the DNA-binding properties of proteins shorter than 259 amino acids, we constructed a set of clones coding for such proteins. The truncated proteins were expressed and purified from both bacterial and mammalian cells followed by examination of their DNA-binding activities.

Various purified proteins isolated from both sources are displayed in Fig. 1A and B. A maximum of four exogenous amino acids can be found at the carboxy terminus of the truncated proteins. These exogenous amino acids are derived from the *Xba*I translational stop linker used in the cloning strategy to generate the expression clones. The exact number of additional amino acids is indicated in the nomenclature of these proteins by the number following the decimal point in their names. The nucleotide and amino acid sequences of the Tag-*Xba*I linker boundary of each protein are shown in Fig. 1C. All the amino-terminal proteins were easily produced in *E. coli* except for those proteins with deletions extending beyond residue 246. Although small quantities of a 242.2-amino-acid protein could be purified, a number of proteins copurified and constituted the majority of the isolated protein (Fig. 1A). Proteins of 240.0 and 235.2 amino acids could not be isolated from *E. coli*, and only small quantities of the 235.2-amino-acid protein could be isolated from HeLa cells (Fig. 1B).

Properties of truncated Tags isolated from *E. coli*. The sequence-specific DNA-binding activity of the purified proteins was examined by a modified McKay assay (43, 68). In this assay, an equimolar mixture of three fragments containing site I, site II, or the wild-type origin (which contains both sites I and II on the same DNA fragment) was used to assess the relative affinity of each protein for these sites in one reaction mixture (Fig. 2D). To ensure that comparisons of origin-specific DNA-binding activity between the truncated proteins were quantifiable, protein concentrations were measured by a dye binding assay and confirmed by electrophoreses on an SDS-polyacrylamide gel.



C

360.1	GAA	CAA	ATG	TTc	tag	tct	aga	cta	g
	Glu	Gln	Met	phe					
249.0	GAA	AGT	TTG	CCc	tag	tct	aga	cta	g
	Glu	Ser	Leu	Pro					
246.4	ATT	GAG	GAA	Act	agt	cta	gac	tag	
	Ile	Glu	Glu	thr	ser	leu	asp		
242.2	CCA	TTT	TCT	cta	gtc	tag	act	ag	
	Pro	Phe	Ser	leu	val				
235.2	TAT	AGT	GCC	cta	gtc	tag	act	ag	
	Tyr	Ser	Ala	leu	val				

FIG. 1. Purified truncated Tags and their sequences in the vicinity of the *Xba*I stop codons. (A) Purified Tags produced in *E. coli* are displayed on an SDS-12.5% polyacrylamide gel stained with Coomassie brilliant blue. Each lane contains 1 μg of protein. The arrow identifies the position of the 242.2-amino-acid protein. (B) Proteins isolated from HeLa cells (1 μg) are shown on a denaturing 15% polyacrylamide gel after silver staining. The arrow identifies the position of the 235.2-amino-acid protein. The numbers on the left indicate molecular size in kilodaltons. (C) The presence of additional amino acids translated from the *Xba*I linker sequence is reflected in the nomenclature of the truncated proteins. Thus, 246.4 refers to a truncated protein containing the N-terminal 246 amino acids from Tag and 4 additional amino acids originating from the *Xba* linker sequences. The 360.1-amino-acid protein was created by insertion of the linker into the *Hpa*I site at nucleotide 3733, while the other truncations were generated by *Bal* 31 treatment. Capital letters represent SV40 nucleotide sequences or amino acid sequences. Lowercase letters represent nucleotide sequences derived from the *Xba*I stop linker or additional amino acids translated from the linker nucleotides. For the sequences of the 272.0-, 266.1-, and 259.3-amino-acid proteins, see reference 68.

The results from Fig. 2A confirm that smaller proteins have greater specific DNA-binding activity (68), culminating in an amino-terminal protein of 246.4 amino acids. To quantitate relative binding affinities between the various proteins, individual bands were excised from the gel (Fig. 2A) and counted in liquid scintillant, and the resultant counts per minute were divided by the number of moles of protein used in the reaction. It was found that proteins containing the amino-terminal 246, 249, 259, and 272 amino acids of Tag when produced in *E. coli* bind to the wild-type origin fragment with relative affinities of 100:50:5:1.7, respectively. Although proteins of 259.3 amino acids and larger were not observed to bind to either site I or site II when less than 1 μg of protein was used in an assay, it was possible to detect binding to these sites with higher quantities of protein (Fig. 2B).

A protein of 242.2 amino acids which lacks only four amino acids from Tag compared with polypeptide 246.4 was found to bind approximately 50-fold less efficiently to the wild-type fragment than polypeptide 246.4 (data not shown). The precise quantitation of relative DNA-binding activity was impossible to determine owing to a lack of adequate amounts of polypeptide 242.2 (Fig. 1A). The very low DNA-binding activity of the 242.2-amino-acid protein and the apparent instability of this and shorter proteins when produced in *E. coli* is consistent with the view that deletions beyond amino acid 246 disrupt a folded protein domain responsible for sequence-specific DNA binding, thus resulting in the production of unstable proteins that are degraded in vivo. A small quantity of a stable 235.2-amino-acid protein was successfully produced in mammalian cells (Fig. 1B). This protein had no detectable DNA-binding activity (data not shown).

Properties of truncated Tags isolated from mammalian cells. SV40 large Tag undergoes substantial posttranslational modifications (7, 23, 29, 36). To investigate what effects these modifications might play in DNA binding, selected clones coding for amino-terminal fragments of Tag were

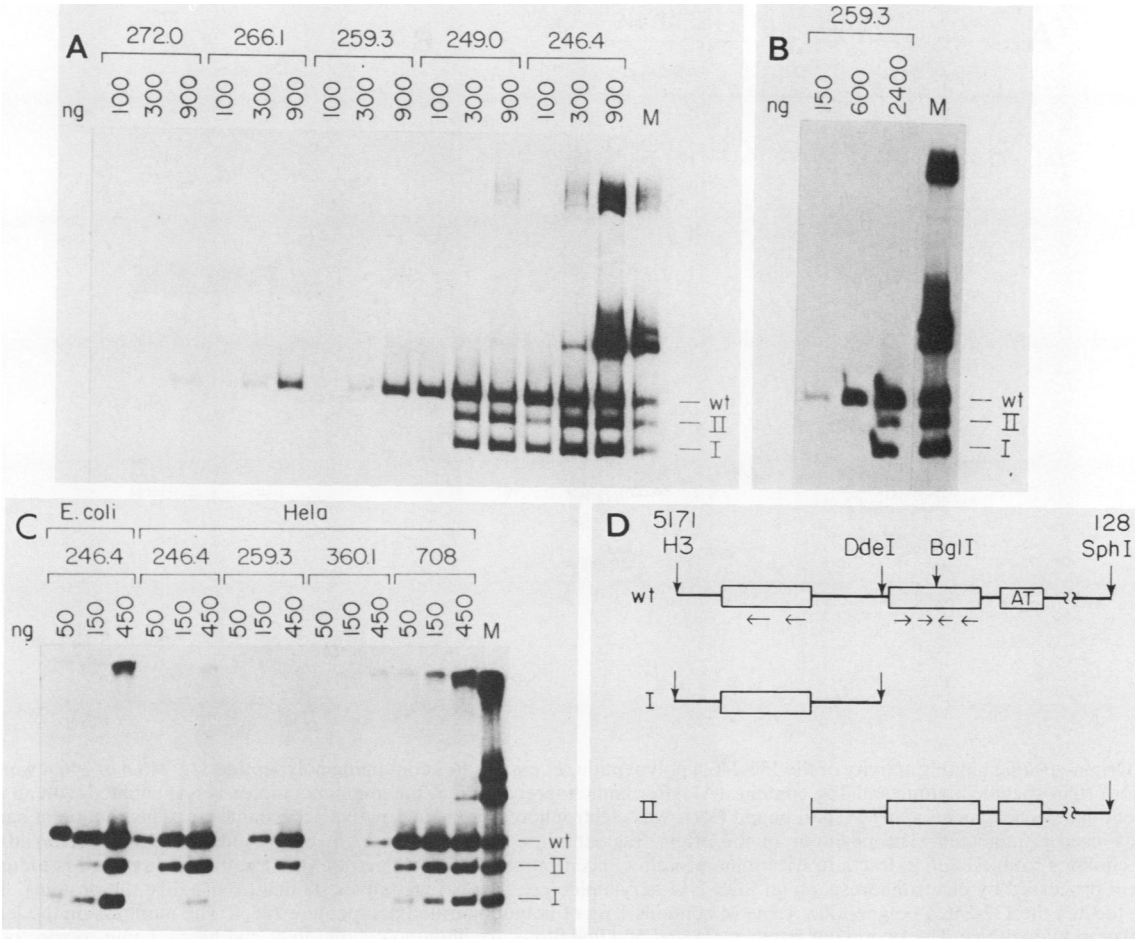


FIG. 2. Origin-specific binding of truncated proteins. Various amounts of purified proteins were incubated with an equimolar mixture of end-labeled fragments (lane M). After 60 min at 0°C, the protein-DNA complex was immunoprecipitated and the bound DNA was electrophoresed on a native 6% polyacrylamide gel. (A and B) Tag proteins purified from *E. coli* assayed for their origin-binding activity. (C) Tag proteins isolated from *E. coli* and HeLa cells assayed for their origin-binding activity. The source, amount, and size of the proteins tested are indicated. (D) DNA templates containing site I, site II or the wild-type (wt) origin from SV40 used in the origin-binding assays. The boxes represent the individual sites, and the horizontal arrows show the pentanucleotide recognition sites of Tag. Vertical arrows show cleavage sites for restriction endonucleases.

expressed in HeLa cells with an adenovirus vector system and the purified proteins displayed in Fig. 1B were assayed for their origin-specific binding properties.

The relative binding activities of these mammalian proteins were calculated as described above. The proteins of 246.4, 259.3, and 360.1 amino acids bound to the wild-type origin DNA fragment with relative ratios of 100:30:3 (Fig. 2C), reaffirming the observation found with the *E. coli* truncated proteins that the smaller proteins had higher DNA-binding activities than the larger proteins.

Mammal- and bacterium-derived truncated proteins were directly compared, and differences relating to their affinities for the wild-type fragment and for the individual sites were observed. When the 246.4-amino-acid protein from both sources was analyzed, the *E. coli*-produced protein bound approximately twofold better to the wild-type fragment than its mammalian counterpart (Fig. 2C). It appears in this case that a protein isolated from HeLa cells is less active than the same protein purified from *E. coli*. However, a comparison of the 259.3-amino-acid protein from the same two sources displayed the opposite effect, with the mammal-derived protein binding three- to fourfold better to the wild-type fragment than the bacterial derivative (extrapolated from data found in Fig. 2A and C and reference 68). A clear distinction was found between the mammalian and bacterial proteins in their preference of binding to the individual site I and site II fragments. The truncated proteins purified from mammalian cells demonstrated a substantially greater affinity for the site II fragment than for the site I fragment (Fig. 2C) (68), while the same proteins isolated from *E. coli* displayed preferential binding to the site I fragment (Fig. 2A, B, and C). The differences in preferential binding between the two sources were well illustrated if the ratio of site II versus site I binding was compared for the 246.4-amino-acid proteins (Fig. 2C). In this comparison, a 10-fold increase in the site II-site I ratio was found between the mammal- and bacterium-derived proteins.

Full-length Tag from HeLa cells was 4- to 10-fold more active in origin-specific DNA binding than the 246.4-amino-acid proteins. There are sequences, therefore, distal to residue 246 necessary for full DNA-binding activity. However, there are also sequences distal to residue 246 that are inhibitory to origin-specific DNA binding since proteins larger than protein 246.4 had reduced levels of activity. The

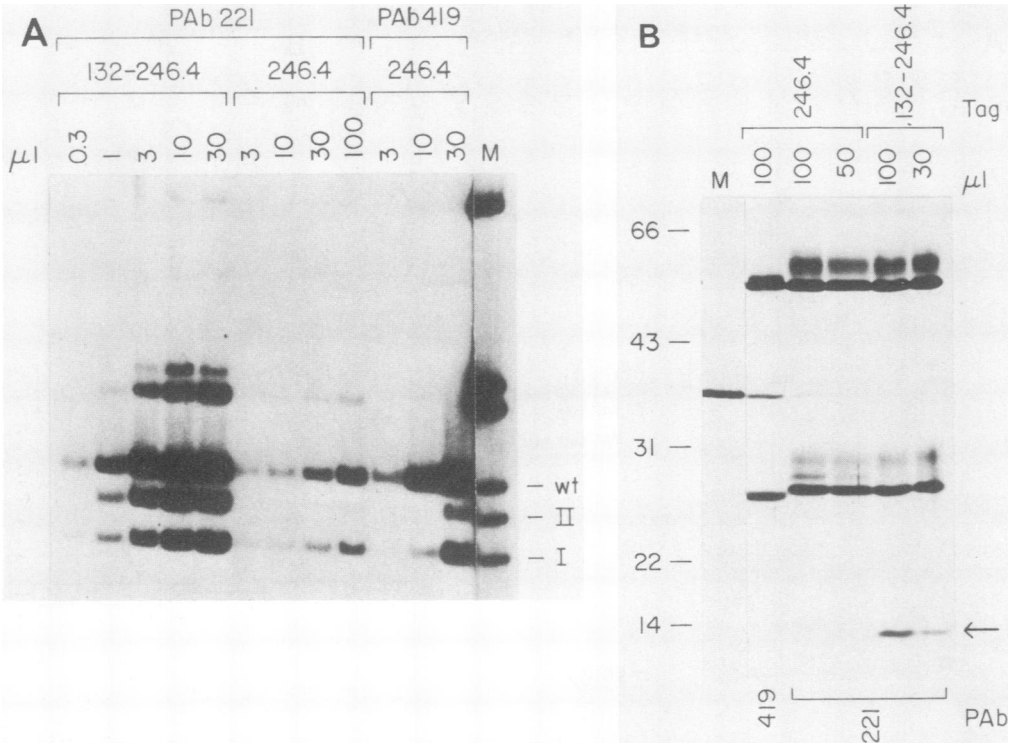


FIG. 3. Origin-specific binding activity of the 132-246.4 polypeptide. *E. coli* lysates containing polypeptide 132-246.4 or 246.4 were assayed in parallel for origin-specific binding and Tag content. (A) After immunoprecipitation, the immunocomplex was incubated with an equimolar mixture of end-labeled fragments (lane M) and bound DNA was electrophoresed on a 6% native acrylamide gel. The volume of each extract and antibody used is indicated. The positions of the origin fragments are shown to the right of the autoradiogram. wt, Wild type. (B) Immunoprecipitates from *E. coli* extracts to determine protein concentrations. The volume of extract and antibody used is indicated. The proteins were processed by electrophoresis on an SDS-17% acrylamide gel. The gel was silver stained to visualize the proteins. The arrow on the right locates the 132-246.4 polypeptide. Lane M contains 1 μg of immunopurified polypeptide 246.4. The numbers on the left indicate molecular size in kilodaltons. The two major bands at 25 and 50 kilodaltons are immunoglobulin light and heavy chain, respectively.

presence of the metal-binding motif between residues 302 and 320 is not sufficient to overcome the inhibitory effect of the sequences distal to residue 246 as evidenced by the low activity of the 360.1-amino-acid protein. Whether this inhibition is biologically significant is not known.

The results obtained with the McKay assay were all reproducible. After multiple repetitions, however, small variations between the different proteins in relative affinities for the wild-type origin DNA fragment were noted. Likewise, the extent of relative affinities displayed by individual proteins for the site I and site II fragments also varied. Nevertheless, the general trends mentioned above were always observed.

DNA-binding domain of SV40 large Tag is located between amino acids 132 and 246. The cumulative data from various laboratories placed the amino-terminal boundary of the origin-specific DNA-binding domain between amino acids 131 (1, 61) and 140 (62). From the work described here, the carboxy-terminal boundary is located at or near amino acid 246. Therefore, it was of interest to determine whether a polypeptide consisting of amino acids 132 to 246 (132-246 polypeptide) would possess origin-specific DNA-binding activity. The polypeptide used in these studies was synthesized in *E. coli* and contained an exogenous initiation methionine and four exogenous amino acids at its carboxy terminus identical to those of the 246.4-amino-acid protein (Fig. 1C). Since this polypeptide lacked the epitope employed to immunopurify the other Tag proteins, our standard purification procedure could not be utilized. Attempts to isolate sufficient amounts of the polypeptide by standard chromatographic techniques were unsuccessful. The origin-specific DNA-binding activity of this polypeptide was therefore determined directly from bacterial extracts by employing a variation of the McKay assay used above. PAb221 (45), a monoclonal antibody that recognizes an epitope between residues 132 and 246 (S. E. Mole and D. P. Lane, personal communication), was used in these experiments. To assess DNA-binding activity of the 132-246.4 protein, the polypeptide was initially immunoprecipitated and washed and then incubated with the identical mixture of origin-bearing DNA fragments used in the previous experiments. This resulted in the 132-246.4 polypeptide binding efficiently in a sequence-specific manner to the wild-type, site I, and site II DNA fragments (Fig. 3A). Thus, this contiguous stretch of 115 amino acids of Tag defines a domain which contains all the structural information necessary to recognize the origin fragments in a sequence-specific manner. Determinants involved in nonspecific DNA binding are also found within this domain.

Relative origin-specific DNA-binding activities between polypeptides 132-246.4 and 246.4 were estimated with proteins from *E. coli* extracts. To compare the DNA-binding activities of these two proteins, PAb419 (27) was required to immunoprecipitate the 246.4-amino-acid protein since PAb221 does not recognize this protein well (Fig. 3B). The reason for the difference in avidity between the two antibod-

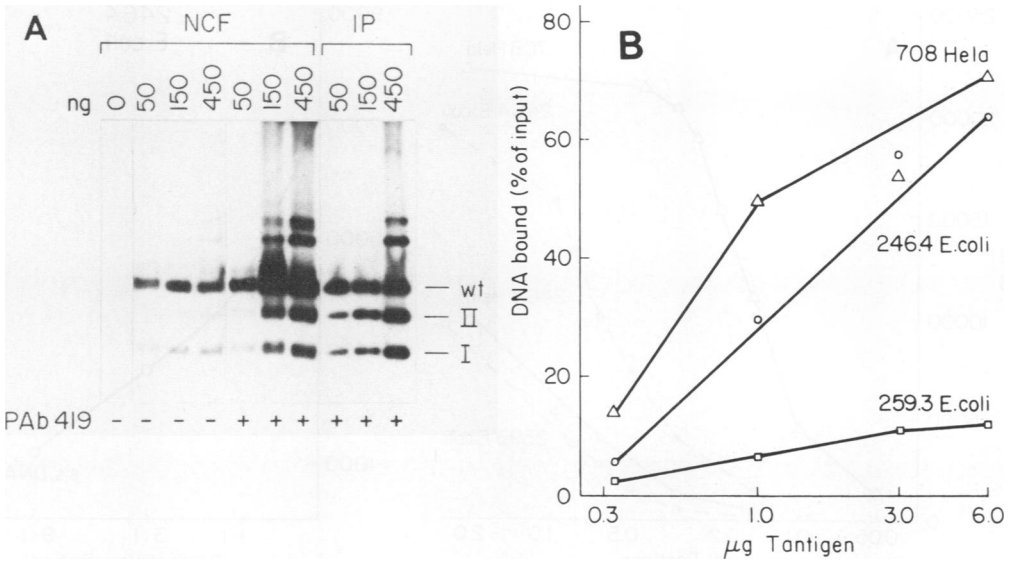


FIG. 4. Origin-specific binding as determined by a nitrocellulose filter binding assay. (A) Increasing amounts of full-length Tag purified from HeLa cells were incubated on ice for 60 min with 5'-labeled DNA fragments (as described in the legend to Fig. 2). Antibody was added to the reaction mixtures as indicated followed by either nitrocellulose filtration (NCF) or immunoprecipitation (IP) of the mixtures. The bound DNA was eluted and subjected to electrophoresis in a 6% native acrylamide gel. The fragments containing the origin sequences are indicated to the right of the autoradiogram. wt, Wild type. (B) Various amounts of truncated and full-length Tag were incubated on ice for 60 min with gel-purified wild-type origin fragment (Fig. 2D). The reaction mixtures were filtered through nitrocellulose and washed, and levels of radioactivity were determined.

ies for polypeptide 246.4 is unknown. By comparing the quantity of protein immunoprecipitated from 100 μl of extract containing protein 132-246.4 by PAb221 and the quantity of protein 246.4 immunoprecipitated by PAb419, it was possible to estimate a twofold-higher concentration of the 132-246.4 polypeptide compared with the 246.4-amino-acid protein in these lysates (Fig. 3B). Since polypeptide 132-246.4 is approximately one-half the size of polypeptide 246.4, this translates into a fourfold molar excess of the shorter protein in the bacterial lysates. Taking this into account, the 1-μl lane of protein 132-246.4 and 3-μl lane of protein 246.4 immunoprecipitated by PAb221 and PAb419, respectively, (Fig. 3A) should contain similar moles of protein. We therefore estimate that both polypeptides have comparable affinities for the wild-type origin fragment, with polypeptide 132-246.4 possibly being slightly more active (twofold). Precise quantitation of sequence-specific DNA-binding activity of polypeptide 132-246.4 will have to await analysis with purified proteins.

Comparison of results obtained by the McKay assay and the nitrocellulose filter binding assay. To confirm the conclusions based on the immunoprecipitation assays used above by an independent method, we used a filter binding assay. Initial experiments were directed at determining the relative affinity that full-length Tag from HeLa cells had for the origin fragments. This was accomplished by incubating Tag for 1 h on ice with an equimolar mixture of origin-bearing DNA fragments identical to those used in the immunoprecipitation experiments. Protein-DNA complexes were subsequently filtered onto nitrocellulose, and the bound DNA was eluted and analyzed by electrophoresis. This experiment was done in both the absence and presence of PAb419, the antibody employed in the McKay assay, to assess the effects this antibody might have on DNA binding (Fig. 4A). In the absence of PAb419, the full-length protein was detected interacting with the site I fragment at lower protein concentrations than that required to detect binding to the site II fragment. The presence of the antibody, however, increased the overall DNA-binding activity of Tag. This enhancement was nearly uniform for all the DNA fragments in the reaction mixture, except that a slightly higher stimulation of site II binding compared with site I binding was observed. The addition of PAb419 to the reaction may influence the filter binding assay by stabilizing the Tag-DNA complex or by increasing the retention of the complex on the nitrocellulose filter. The results obtained with the nitrocellulose filter binding assay in the presence of PAb419 were very similar to those of the McKay assay (Fig. 4A). Tag was detected preferentially complexing with the wild-type fragment versus the site I and site II fragments. Although Tag displayed approximately equal affinity for sites I and II in the McKay assay (Fig. 4A), the protein was predominantly observed complexing preferentially to site I rather than site II (data not shown) (68). However, a slight preference by Tag for the site II fragment was also observed (Fig. 2C) (7). Nevertheless, the overall results obtained from the McKay and nitrocellulose filter binding assays were very similar.

The nitrocellulose filter binding assay was further employed to assess relative affinities for the wild-type origin fragment between full-length Tag purified from mammalian cells and the amino-terminal proteins 259.3 and 246.4 isolated from *E. coli*. This was accomplished by titrating each protein against a constant amount of wild-type fragment that had been gel purified. The shorter amino-terminal protein was more active than the larger truncated protein, while the full-length protein contained the highest origin-specific DNA-binding activity of the three proteins (Fig. 4B). The same hierarchy of relative DNA-binding activities between these three proteins was also observed when the McKay assay was employed. This suggests that the McKay and nitrocellulose filter binding assays are detecting the same general Tag-DNA complex formation. The methodologies,

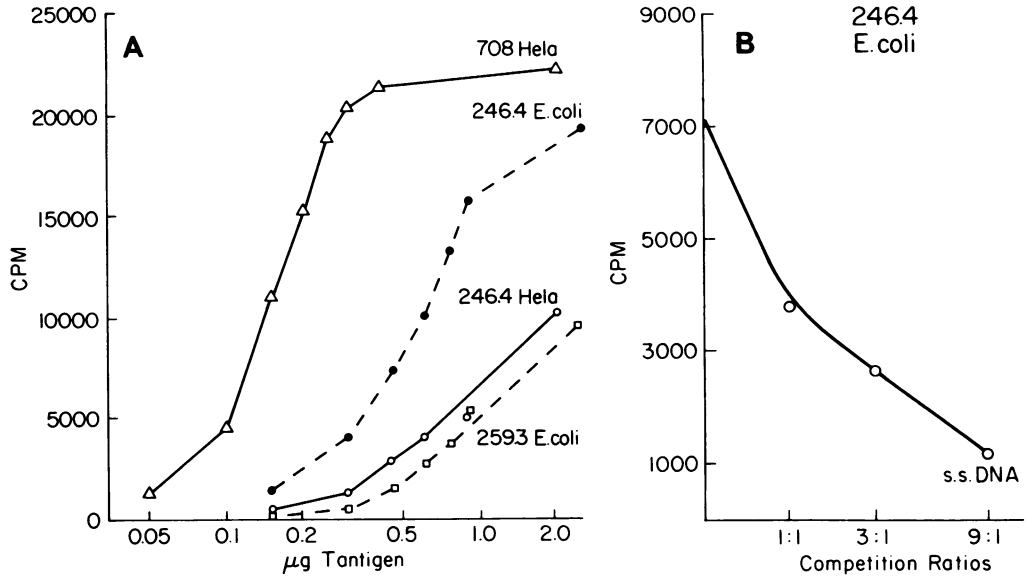


FIG. 5. Single-stranded DNA-binding activities of full-length and truncated Tags. (A) A heteroduplex substrate consisting of a single-stranded circular DNA annealed to a ^{32}P-labeled oligonucleotide was incubated with various amounts of protein for 60 min on ice. The reaction mixtures were filtered through nitrocellulose, washed, and assayed for levels of radioactivity. (B) The ability of polypeptide 246.4 from *E. coli* to complex with the labeled heteroduplex substrate, employing nonsaturating amounts of protein, was decreased in a linear fashion by increasing amounts of nonlabeled single-stranded DNA. Protein (500 ng) and 20 ng of heteroduplex along with either 0, 20, 60, or 180 ng of single-stranded DNA were incubated on ice for 60 min. The reaction mixtures were filtered through nitrocellulose and washed, and levels of radioactivity were determined.

however, are not identical since relative activity levels among the three proteins are different for the two assays. The differences recorded between the two methodologies could be attributed to the antibody-antigen interaction of the McKay assay preferentially stabilizing the DNA-binding activity of the shorter proteins or possibly to differential sensitivity of the protein-DNA complexes to denaturation on the nitrocellulose filter.

Single-stranded DNA-binding activity is associated with the amino-terminal 246.4-amino-acid protein. In contrast to the wealth of data defining the origin-specific DNA-binding domain, the precise region of Tag involved in binding single-stranded DNA has not been identified. To evaluate the single-stranded DNA-binding affinities of the various proteins, we used a heteroduplex helicase substrate in conjunction with a nitrocellulose filter binding assay. Each protein was incubated for 1 h on ice with circular single-stranded DNA (mpSV2) annealed with an α-^{32}P-labeled oligonucleotide. Protein-DNA complexes were subsequently filtered onto nitrocellulose, and levels of radioactivity were determined. The single-stranded DNA-binding domain was localized to the amino-terminal 246 amino acids of Tag (Fig. 5A). To ensure that single-stranded DNA binding was being assayed, competition between the labeled heteroduplex and single-stranded DNA was assayed. When either full-length Tag from mammalian cells (data not shown) or the amino-terminal 246.4-amino-acid protein isolated from *E. coli* (Fig. 5B) was tested in the competition assay, detection of heteroduplex binding was decreased in a 1:1 molar ratio by single-stranded DNA, thus confirming that single-stranded DNA binding was being assayed. As was found with origin-specific binding, the metal-binding motif located between residues 302 and 320 is dispensable for single-stranded DNA-binding activity.

By comparing the single-stranded DNA-binding activities of polypeptides 246.4 and 259.3 produced in *E. coli*, one can see that the shorter protein is detected complexing with single-stranded DNA more efficiently, recapitulating the same phenomenon observed in the origin-specific DNA-binding assays. However, wild-type levels of single-stranded DNA-binding activity require carboxy sequences missing from the truncated proteins as evidenced by full-length Tag derived from HeLa cells. To investigate the possibility that the source of the protein affected single-stranded DNA-binding activity, we compared the amino-terminal 246.4-amino-acid proteins derived from mammalian and bacterial cells. The truncated protein isolated from *E. coli* was detected complexing with single-stranded DNA at lower protein concentrations than its mammalian counterpart, implicating posttranslational modifications affecting single-stranded DNA-binding activity.

DISCUSSION

To better understand the biological functions of Tag, we defined the protein domain responsible for DNA binding. Tag binds specifically to two unique sites in the SV40 origin of replication (sites I and II) and to single-stranded DNA. These three DNA-binding activities colocalize to the amino-terminal 246 amino acids of the protein. The polypeptide consisting of residues 132 to 246 is an independent domain capable of binding efficiently to both sites in the SV40 origin in a sequence-specific manner. We propose that this domain is also responsible for the single-stranded DNA binding of Tag. All three of these DNA-binding activities appear to be regulated by posttranslational modifications, because amino-terminal fragments of Tag when produced in mammalian and bacterial cells had distinct DNA-binding properties. Interestingly, as the amino-terminal proteins decreased in size, their levels of origin-specific DNA-binding activity increased from near background to levels approaching those of the full-length protein.

To localize the protein domain responsible for single-stranded and origin-specific DNA-binding activity of Tag, a series of progressively smaller amino-terminal proteins were purified to near homogeneity and their DNA-binding activity was determined. As the truncated proteins declined in size, their ability to bind single-stranded and origin-specific DNA increased progressively. This held true for both mammal- and bacterium-derived proteins. The larger, less active truncated proteins could be detected binding to the SV40 origin-specific sequences only when large quantities of purified protein were assayed. When these proteins were tested directly from cellular extracts, however, they displayed no detectable origin-specific DNA-binding activity (10, 48), potentially giving the DNA-binding domain a false carboxy-terminal boundary. In addition, a sharp peak in activity spanning approximately four amino acids (246 to 249) was found with truncated proteins that define the carboxy boundary of the DNA-binding domain. Proteins either slightly larger or smaller (polypeptides 259.3 and 242.2, respectively), separated by only 17 amino acids, were a minimum of 10-fold less active in origin-specific DNA binding. Identifying the sequence-specific DNA-binding domain of other proteins may also require the use of purified proteins of several closely spaced deletion mutants. This will ensure that a false boundary of the DNA-binding domain of the protein is not merely confirmed by a large deletion mutant that extends into the DNA-binding domain. The resurgence of DNA-binding activity by the shorter proteins of Tag continued to increase until a maximum of activity was found in a protein of 246.4 amino acids. Sequences inhibitory to DNA binding apparently begin very near residue 246 since deletions down to this amino acid continued to have increasing levels of DNA-binding activity. Mutations C terminal to amino acid 246, which include mutants in the putative metal-binding motif (1, 39) as well as the temperature-sensitive mutants (73), also have reduced origin-specific DNA-binding activity. The sequence-specific DNA-binding activity of two transcription factors, Sp1 (31, 32) and even-skipped (28), are similarly influenced by protein sequences distal to their DNA-binding domain. Although sequences distal to amino acid 246 inhibit origin-specific and single-stranded DNA-binding activities, the full-length protein from mammalian cells exhibited higher DNA-binding activities for both these properties. The enhanced binding activity displayed by full-length Tag may reflect an ability to contour the DNA-binding domain into an optimal configuration not obtainable by the smaller proteins. Likewise, it is possible that the truncated proteins are defective in an ancillary function required to stabilize protein-DNA interactions. Oligomerization is known to stabilize protein-DNA complexes (46, 54). Full-length Tag forms a multimeric nucleoprotein complex with the SV40 origin of replication (5, 12, 15) as well as oligomeric structures in solution (19). The truncated proteins, however, are found as monomers in solution (I. Mohr, personal communication). It is, therefore, conceivable that the defect in oligomerization of the truncated protein is responsible for its reduced DNA-binding activity.

The single-stranded DNA-binding activity of Tag may reside in the same protein domain responsible for origin-specific DNA binding. Both activities localized to the first 246 amino acids of Tag and behaved in an analogous fashion with regards to the truncated and full-length protein as just described. The abilities of polypeptide 246.4 and full-length Tag to bind single-stranded DNA were competed against in qualitatively the same manner by DNA homopolymers, confirming the single-stranded DNA binding of polypeptide 246.4 is a specific event (data not shown). The amino-terminal boundary of the single-stranded DNA-binding domain can be assigned to residue 115 or larger since a deletion mutant of Tag (D2) which lacks the amino-terminal 114 amino acids is capable of unwinding a linear double-stranded DNA fragment (34). This activity requires Tag to bind single-stranded DNA. The hypothesis that single-stranded and origin-specific DNA binding reside in the same domain is further bolstered by the observation that a number of aromatic and basic residues found within polypeptide 132-246 are conserved among papovavirus large Tag. In the well-characterized single-stranded DNA-binding proteins encoded by bacteriophage Fd gene 5 and T4 gene 32, aromatic residues have been shown to interact with the bases either by intercalating between the bases of the nucleotides or associating with them in a hydrophobic pocket (11). The basic residues of these proteins are thought to interact with the DNA sugar-phosphate backbone (11). Although the DNA-binding activities of Tag may co-localize to the same physical domain in the protein, the determinants which specify single-stranded and origin-specific DNA binding appear to be quite different. This notion is supported by mutations within the origin-specific DNA-binding domain which render Tag incapable of sequence-specific DNA binding but display wild-type levels of single-stranded DNA binding (3, 44a). Furthermore, deletion mutants of Tag are capable of separating their ability to bind to either site I or site II. However, the interaction by Tag with the two origin sequences must be closely linked since point mutations inside the DNA-binding domain that disrupt origin-specific binding activity abolish both site I and site II binding (44a). These data imply that Tag exists in three unique conformations, each responsible for single-stranded, site I, or site II DNA binding.

The expression of a mammalian gene product in an *E. coli* system most probably results in a protein lacking mammalian posttranslational modifications, and as such, these proteins may have different properties than proteins purified from mammalian cells. Thus, any conclusions ascribed to a mammalian protein derived from bacterial sources must be confirmed by studies of the mammalian derivative of the protein. In the case of full-length Tag, the abilities of the *E. coli*-produced protein to bind to site II and function in a replication assay were dramatically reduced, while its ATPase and helicase activities were unaffected compared with those of the HeLa-derived protein (44b). As with full-length Tag, the truncated proteins from mammalian and bacterial cells retained a variety of basic functions but diverged in specificity and levels of activity. Amino-terminal proteins of 259.3 and 266.1 amino acids bound with an approximate threefold-higher affinity for the wild-type origin fragment when produced in mammalian cells than they did when produced in bacterial cells (Fig. 2) (68). The specificity of DNA binding is also affected by the source of the protein, which is most apparent when comparing polypeptide 246.4 from mammalian and bacterial cells. The bacterial derivative of this protein bound efficiently to both site I and II, while the mammalian derivative bound efficiently to site II with little or no detectable site I binding. These differences in activities strongly suggest that mammalian posttranslational modifications alter the content of the DNA-binding domain and thus alter its level of activity and specificity for the individual origin sites. The inability of mammalian posttranslational modifications to enhance the DNA-binding activity of polypeptide 246.4 is consistent with the notion that the DNA-binding domain in this protein is in an optimal config-

uration. This idea is supported by the observation that protein 246.4 contains the highest DNA-binding activity of all the truncated proteins. The source of the protein also proved to have an effect on single-stranded DNA-binding activity, with the *E. coli*-produced protein binding better than its mammalian counterpart.

Phosphorylation is the posttranslational modification most likely responsible for the differences in DNA-binding activities between the mammal- and bacterium-derived truncated proteins. Biochemical and genetic analysis has demonstrated posttranslational phosphorylation of Tag to be a modulator of origin-specific DNA-binding activity (35, 44, 63). Tag is phosphorylated at two clusters of amino acids, one at the carboxy terminus and the other at the amino terminus of the protein (55); both lie outside the DNA-binding domain. Each cluster contains four serines and one threonine as sites of phosphorylation. The amino-terminal truncated proteins retain the amino-terminal phosphorylation cluster. Biochemical analysis has shown that the phosphoserines on Tag are susceptible to dephosphorylation by calf intestinal alkaline phosphatase, while the phosphothreonines are refactory (35, 58). Calf intestinal alkaline phosphatase treatment of full-length mammalian Tag stimulates its DNA-binding activity (35, 44, 63). However, calf intestinal alkaline phosphatase treatment of the mammal- and bacterium-derived truncated proteins had no effect on their ability to bind DNA, suggesting that the amino-terminal serines require the entire protein to exert their effect or are not involved in regulating DNA binding (68). Genetic analysis supports this latter notion, as single mutations which convert the amino-terminal phosphoserine residues into alanines do not affect the DNA-binding properties of the full-length protein (57). In striking contrast, when the amino-terminal threonine at residue 124 is mutated to an alanine, site II binding is completely lost while binding to the wild-type origin is maintained, suggesting that site I binding is maintained (57). Some mutations of the carboxy-terminal phosphoamino acid residues had little or a much reduced effect on site II binding of Tag, although a mutation at serine 677 did affect site I binding (57). This is opposite to the mutations of threonine 124. These data strongly suggest that phosphorylation of threonine 124 is the causative agent for the differences in activities between the mammal- and bacterium-produced truncated proteins.

Tag is a model system to study both static and dynamic nucleoprotein complexes. However, the structural conformation of the DNA-binding domain from Tag has yet to be elucidated. The helix-turn-helix motif found in various bacterial DNA-binding proteins (54, 66) and in eucaryotic homeobox DNA-binding proteins (37, 47) is missing from Tag. Likewise, the zinc finger required for DNA binding in various proteins (18, 25) is dispensable for both origin-specific (68) and single-stranded DNA binding. It is possible that the physical requirements placed on the DNA-binding domain of Tag for it to interact with single-stranded DNA and two unique sites in the viral origin of replication obscure a simple motif involved in complexing with DNA, or a novel motif may be required for such complexity. Having defined the protein domain responsible for single-stranded and origin-specific DNA binding, it is now possible to study its structure and regulation with greater clarity.

ACKNOWLEDGMENTS

We are grateful to Sara Mole and David Lane for the generous gift of PAb221. We thank Bruce Stillman, Mike Gilman, Leonardo Brizuela, Winship Herr, and David Beach for critical reading of the manuscript. We most heartily thank Ian Moh for his helpful discussions and suggestions throughout this endeavor. We thank Marilyn Goodwin for her excellent help in preparing the manuscript.

This work was funded by Public Health Service grant CA 13106 from the National Cancer Institute to Cold Spring Harbor Laboratory.

ADDENDUM

The 259.3-amino-acid protein derived from bacterial cells can be phosphorylated in vitro uniquely on threonine 124 by employing the human p34-cdc2 kinase. This phosphorylation event stimulated sequence-specific binding to the wild-type and site II origin fragments 20-fold, while site I binding was unaffected. Similarly, a 10-fold stimulation in the site II binding activity of full-length Tag from bacteria occurred after treatment with $p34^{cdc2}$ kinase. Furthermore, the phosphorylated protein was more active than the mammalian protein in an in vitro replication assay (D. McVey, L. Brizuela, I. J. Mohr, D. R. Marshak, Y. Gluzman, and D. Beach, Nature [London] **341**:503–507).

LITERATURE CITED

1. **Arthur, A. K., A. Hoss, and E. Fanning.** 1988. Expression of simian virus 40 T antigen in *Escherichia coli*: localization of T-antigen origin DNA-binding domain to within 129 amino acids. J. Virol. **62**:1999–2006.
2. **Auborn, K., M. Guo, and C. Prives.** 1989. Helicase, DNA binding, and immunological properties of replication-defective simian virus 40 mutant T antigens. J. Virol. **63**:912–918.
3. **Auborn, K. J., R. B. Markowitz, E. Wang, Y. T. Yu, and C. Prives.** 1988. Simian virus (SV40) T antigen binds specifically to double-stranded DNA but not to single-stranded DNA or DNA-RNA hybrids containing the SV40 regulatory sequences. J. Virol. **62**:2204–2208.
4. **Berg, J. M.** 1986. Potential metal binding domains in nucleic acid binding proteins. Science **232**:485–487.
5. **Borowiec, J. A., and J. Hurwitz.** 1988. ATP stimulates the binding of simian virus 40 (SV40) large tumor antigen to the SV40 origin of replication. Proc. Natl. Acad. Sci. USA **85**: 64–68.
6. **Borowiec, J. A., and J. Hurwitz.** 1988. Localized melting and structural changes in the SV40 origin of replication induced by T antigen. EMBO J. **7**:3149–3158.
7. **Bradley, M. K., J. Hudson, M. S. Villanueva, and D. M. Livingston.** 1984. Specific in vitro adenylation of the simian virus 40 large tumor antigen. Proc. Natl. Acad. Sci. USA **81**:6574–6578.
8. **Bradley, M. K., T. F. Smith, R. H. Lathrop, D. M. Livingston, and T. A. Webster.** 1987. Consensus topography in the ATP binding site of simian virus 40 and polyomavirus large tumor antigens. Proc. Natl. Acad. Sci. USA **84**:4026–4030.
9. **Burger, C., and E. Fanning.** 1983. Specific DNA binding activity of T antigen subclasses varies among different SV40-transformed cell lines. Virology **126**:19–31.
10. **Clark, R., K. Peden, J. M. Pipas, D. Nathans, and R. Tjian.** 1983. Biochemical activities of T-antigen proteins encoded by simian virus 40 *A* gene deletion mutants. Mol. Cell. Biol. **3**:220–228.
11. **Coleman, J. E., K. R. Williams, G. C. King, R. V. Prigodich, Y. Shamoo, and W. H. Konigsberg.** 1987. Mapping the functional domains in single strand DNA binding proteins gene 5 and gene 32, p. 323–336. In Dale L. Oxender and C. F. Fox (ed.), Protein engineering. Alan R. Liss, Inc., New York.
12. **Dean, F. B., M. Dodson, H. Echols, and J. Hurwitz.** 1987. ATP dependent formation of a specialized nucleoprotein structure by simian virus 40 (SV40) large tumor antigen at the SV40 replication origin. Proc. Natl. Acad. Sci. USA **84**:8981–8985.
13. **Deb, S., A. L. DeLucia, C. P. Baur, A. Koff, and P. Tegtmeyer.** 1986. Domain structure of the simian virus 40 core origin of replication. Mol. Cell. Biol. **6**:1663–1670.
14. **Deb, S., A. L. DeLucia, A. Koff, S. Tsui, and P. Tegtmeyer.** 1986. The adenine-thymine domain of the simian virus 40 core origin

directs DNA bending and coordinately regulates DNA replication. Mol. Cell. Biol. 6:4578–4584.
15. **Deb, S. P., and P. Tegtmeyer.** 1987. ATP enhances the binding of simian virus 40 large T antigen to the origin of replication. J. Virol. 61:3649–3654.
16. **DeLucia, A. L., B. A. Lewton, R. Tjian, and P. Tegtmeyer.** 1983. Topography of simian virus 40 A protein-DNA complex: arrangement of pentanucleotide interaction sites at the origin of replication. J. Virol. 46:143–150.
17. **Dixon, R. A. F., and D. Nathans.** 1985. Purification of simian virus 40 large T antigen by immunoaffinity chromatography. J. Virol. 53:1001–1004.
18. **Evans, R. M., and S. M. Hollenberg.** 1988. Zinc fingers: gilt by association. Cell 52:1–3.
19. **Fanning, E., K. H. Westphal, D. Brauer, and D. Corlin.** 1982. Subclasses of simian virus 40 large T antigen: differential binding of two subclasses of T antigen from productively infected cells to viral and cellular DNA. EMBO J. 11:1023–1028.
20. **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.
21. **Gish, W. R., and M. R. Botchan.** 1987. Simian virus 40 transformed human cells that express large T antigens defective for viral DNA replication. J. Virol. 61:2864–2876.
22. **Gluzman, Y., R. Frisque, and J. Sambrook.** 1980. Origin defective mutants of SV40. Cold Spring Harbor Symp. Quant. Biol. 44:293–298.
23. **Goldman, N., M. Brown, and G. Khoury.** 1981. Modification of SV40 T antigen by poly ADP ribosylation. Cell 24:567–572.
24. **Grasser, F. A., K. Mann, and G. Walter.** 1987. Removal of serine phosphates from simian virus 40 large T antigen increases its ability to stimulate DNA replication in vitro but has no effect on ATPase and DNA binding. J. Virol. 61:3373–3380.
25. **Hanas, J. S., D. J. Hazuda, D. F. Bogenhagen, F. Y. H. Wu, and C. W. Wu.** 1983. Xenopus transcription factor IIIA requires zinc for binding to the 5S RNA gene. J. Biol. Chem. 258: 14120–14125.
26. **Hanson, U., D. G. Tenen, D. M. Livingston, and P. A. Sharp.** 1981. T antigen repression of SV40 early transcription from two promoters. Cell 27:603–612.
27. **Harlow, E., L. V. Crawford, D. C. Pim, and N. M. Williamson.** 1981. Monoclonal antibodies specific for simian virus 40 tumor antigens. J. Virol. 39:861–869.
28. **Hoey, T., R. Warrior, J. Manak, and M. Levine.** 1988. DNA-binding activities of the *Drosophila melanogaster* even-skipped protein are mediated by its homeo domain and influenced by protein context. Mol. Cell. Biol. 8:4598–4607.
29. **Jarvis, D. L., and J. S. Butel.** 1985. Modification of simian virus 40 large tumor antigen by glycosylation. Virology 141:173–189.
30. **Jones, K. A., and R. Tjian.** 1984. Essential contact residues within SV40 large T antigen binding sites I and II identified by alkylation interference. Cell 36:155–162.
31. **Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. T. Tjian.** 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell 51:1079–1090.
32. **Kadonaga, J. T., A. J. Courey, J. Ladika, and R. Tjian.** 1988. Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science 242:1566–1569.
33. **Kalderon, D., and A. E. Smith.** 1984. In vitro mutagenesis of a putative DNA binding domain of SV40 large T antigen. Virology 139:109–137.
34. **Klausing, K., M. Scheffner, K. H. Scheidtmann, H. Stahl, and R. Knippers.** 1989. DNA binding properties and replication activity of the T antigen related D2 phosphoprotein. Biochemistry 28:2238–2244.
35. **Klausing, K., K. H. Scheidtmann, E. A. Baumann, and R. Knippers.** 1988. Effects of in vitro dephosphorylation on DNA binding and DNA helicase activities of simian virus 40 large tumor antigen. J. Virol. 62:1258–1265.
36. **Klockmann, Y., and W. Deppert.** 1983. Acylated simian virus 40 large T antigen: a new subclass associated with a detergent resistant lamina of the plasma membrane. EMBO J. 2:1151–1157.
37. **Laughon, A., and M. P. Scott.** 1984. Sequence of a *Drosophila* segmentation gene protein structure homology with DNA-binding proteins. Nature (London) 310:25–31.
38. **Li, J. J., K. W. C. Peden, R. A. F. Dixon, and T. Kelly.** 1986. Functional organization of the simian virus 40 origin of DNA replication. Mol. Cell. Biol. 6:1117–1128.
39. **Loeber, G., R. Parsons, and P. Tegtmeyer.** 1989. The zinc finger region of simian virus 40 large T antigen. J. Virol. 63:94–100.
40. **Margolskee, R., and D. Nathans.** 1984. Simian virus 40 mutant T antigens with relaxed specificity for the nucleotide sequence at the viral DNA origin of replication. J. Virol. 49:386–393.
41. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.
42. **McEntee, K., G. M. Weinstock, and I. R. Lehman.** 1980. recA protein-catalyzed strand assimilation: stimulation by *Escherichia coli* single-stranded DNA-binding protein. Proc. Natl. Acad. Sci. USA 77:857–861.
43. **McKay, R. D. G.** 1981. Binding of a simian virus 40 large T antigen related protein to DNA. J. Mol. Biol. 145:471–488.
44. **Mohr, I. J., B. Stillman, and Y. Gluzman.** 1987. Regulation of SV40 DNA replication by phosphorylation of T antigen. EMBO J. 6:153–160.
44a. **Mohr, I. J., M. P. Fairman, B. Stillman, and Y. Gluzman.** 1989. Large T-antigen mutants define multiple steps in the initiation of simian virus 40 DNA replication. J. Virol. 63:4181–4188.
44b. **Mohr, I. J., Y. Gluzman, M. P. Fairman, M. Strauss, D. McVey, B. Stillman, and R. Gerard.** 1989. Production of simian virus 40 large tumor antigen in bacteria: altered DNA-binding specificity and DNA-replication activity of underphosphorylated large tumor antigen. Proc. Natl. Acad. Sci. USA 86:6479–6483.
45. **Mole, S. E., J. V. Gannon, M. J. Ford, and D. P. Lane.** 1987. Structure and function of SV40 large T antigen. Philos. Trans. R. Soc. London Ser. B 317:455–469.
46. **Newport, J. W., N. Longerg, S. C. Kowalckowski, and P. H. von Hippel.** 1981. Interactions of bacteriophage T4 gene 32 protein with nucleic acids. J. Mol. Biol. 145:105–121.
47. **Otting, G., Y. Qian, M. Muller, M. Affolter, W. Gehring, and D. Wutrich.** 1988. Secondary structure determination for the *Antennapedia* homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. EMBO J. 7:4305–4309.
48. **Paucha, E., D. Kalderon, R. W. Harvey, and A. E. Smith.** 1986. Simian virus 40 origin DNA-binding domain on large T antigen. J. Virol. 57:50–64.
49. **Prives, C., L. Covey, A. Scheller, and Y. Gluzman.** 1983. DNA-binding properties of simian virus 40 T-antigen mutants defective in viral DNA replication. Mol. Cell. Biol. 3:1958–1966.
50. **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. USA 77:5706–5710.
51. **Rio, D. C., and R. Tjian.** 1983. SV40 T antigen binding site mutations that affect autoregulation. Cell 32:1227–1240.
52. **Ryder, K., S. Silver, A. L. DeLucia, E. Fanning, and P. Tegtmeyer.** 1986. An altered DNA conformation in origin region I is a determinant for the binding of SV40 large T antigen. Cell 44:719–725.
53. **Ryder, K., E. Vakalopoulou, R. Mertz, I. Mastrangelo, P. Hough, P. Tegtmeyer, and E. Fanning.** 1986. Seventeen base pairs of region I encode a novel tripartite binding signal for SV40 T antigen. Cell 42:539–548.
54. **Sauer, R. T., R. R. Yocum, R. F. Doolittle, M. Lewis, and C. O. Pabo.** 1982. Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. Nature (London) 298:447–451.
55. **Scheidtmann, K. H., B. Echle, and G. Walter.** 1982. Simian virus 40 large T antigen is phosphorylated at multiple sites contained in two separate regions. J. Virol. 44:116–133.
56. **Scheidtmann, K. H., M. Hardung, B. Echle, and G. Walter.** 1984. DNA-binding activity of simian virus 40 large T antigen correlates with a distinct phosphorylation state. J. Virol. 50: 1–12.
57. **Schneider, J., and E. Fanning.** 1988. Mutations in the phosphor-

ylation sites of simian virus 40 (SV40) T antigen alter its origin DNA-binding specificity for site I or II and affect SV40 DNA replication activity. J. Virol. **62**:1598–1605.
58. **Shaw, S. B., and P. Tegtmeyer.** 1981. Binding of dephosphorylated A protein to SV40 DNA. Virology **115**:88–96.
59. **Shortle, D. A., 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. USA **76**:6128–6131.
60. **Simanis, V., and D. P. Lane.** 1985. An immunoaffinity purification for SV40 large T antigen. Virology **4**:88–100.
61. **Simmons, D. T.** 1986. DNA-binding region of the simian virus 40 tumor antigen. J. Virol. **57**:776–785.
62. **Simmons, D. T.** 1988. Geometry of the simian virus 40 large tumor antigen-DNA complex as probed by protease digestion. Proc. Natl. Acad. Sci. USA **85**:2086–2090.
63. **Simmons, D. T., W. Chou, and K. Rodgers.** 1986. Phosphorylation downregulates the DNA-binding activity of simian virus 40 T antigen. J. Virol. **60**:888–894.
64. **Spillman, T., D. Giacherio, and L. P. Hager.** 1979. Single strand DNA binding of simian virus 40 tumor antigen. J. Biol. Chem. **254**:3100–3104.
65. **Stahl, H., P. Droge, and R. Knippers.** 1986. DNA helicase activity of SV40 large tumor antigen. EMBO J. **5**:1939–1944.
66. **Steitz, T. A., D. H. Ohlendort, D. B. McKay, W. F. Anderson, and B. W. Mathews.** 1982. Structural similarities in the DNA-binding domains of catabolite gene activator and Cro repressor protein. Proc. Natl. Acad. Sci. USA **79**:3097–3100.
67. **Stillman, B., R. D. Gerard, R. A. Guggenheimer, and Y. Gluzman.** 1985. T antigen and template requirements for SV40 DNA replication in vitro. EMBO J. **4**:2933–2939.
68. **Strauss, M., P. Argani, I. J. Mohr, and Y. Gluzman.** 1987. Studies on the origin-specific DNA-binding domain of simian virus 40 large T antigen. J. Virol. **61**:3326–3330.
69. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. J. Mol. Biol. **189**:113–130.
70. **Tegtmeyer, P.** 1972. Simian virus 40 deoxyribonucleic acid synthesis: the viral replicon. J. Virol. **10**:591–598.
71. **Tegtmeyer, P., M. Schwartz, J. K. Collins, and K. Rundell.** 1975. Regulation of tumor antigen synthesis by simian virus 40 gene A. J. Virol. **16**:168–178.
72. **Tooze, J. (ed.)** 1981. Molecular biology of tumor viruses, part 2. DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
73. **Wilson, V. G., M. J. Tevethia, B. A. Lewton, and P. Tegtmeyer.** 1982. DNA-binding properties of simian virus 40 temperature-sensitive A proteins. J. Virol. **44**:458–466.
74. **Wold, M. S., J. J. Li, and T. J. Kelly.** 1987. Initiation of simian virus 40 DNA replication in vitro: large tumor antigen and origin dependent unwinding of the template. Proc. Natl. Acad. Sci. USA **84**:3643–3647.