JOURNAL OF VIROLOGY, Oct. 1990, p. 4858–4865
0022-538X/90/104858-08$02.00/0

Vol. 64, No. 10

Identification of Simian Virus 40 T-Antigen Residues Important for Specific and Nonspecific Binding to DNA and for Helicase Activity

DANIEL T. SIMMONS,* KYUNGOK WUN-KIM, AND WILLIAM YOUNG
School of Life and Health Sciences, University of Delaware, Newark, Delaware 19716

Received 9 April 1990/Accepted 6 July 1990

We have previously identified three regions (called elements) in the DNA-binding domain of simian virus 40 large tumor (T) antigen which are critical for binding of the protein to the recognition pentanucleotides GAGGC at the viral replication origin. These are elements A (residues 147 to 159), B1 (185 to 187), and B2 (203 to 207). In this study, we generated mutants of simian virus 40 in order to make single-point substitution mutations at nearly every site in these three elements. Each mutation was tested for its effect on virus replication, and T antigen was produced from all replication-negative mutants. The mutant proteins were assayed for binding to several different DNA substrates and for helicase activity. We found that within each element, mutations at some sites had major effects on DNA binding while mutations at other sites had moderate, mild, or minimal effects, suggesting that some residues are more important than others in mediating DNA binding. Furthermore, we provide evidence that certain residues in elements A and B2 (Ala-149, Phe-159, and His-203) participate in nonspecific double-stranded and helicase substrate (single-stranded) DNA binding while others (Ser-147, Ser-152, Asn-153, Thr-155, Arg-204, Val-205, and Ala-207) are involved in sequence-specific binding at the origin. The residues in element B1 (primarily Ser-185 and His-187) take part only in nonspecific DNA binding. The amino acids important for nonspecific DNA binding are also required for helicase activity, and we hypothesize that they make contact with the sugar-phosphate backbone of DNA. On the other hand, those involved in sequence-specific binding are not needed for helicase activity. Finally, our analysis showed that three residues (Asn-153 and Thr-155 in element A and Arg-204 in element B2) may be the most important for sequence-specific binding. They are likely to make direct or indirect contacts with the pentanucleotide sequences at the origin.

The mechanism by which simian virus 40 (SV40) large tumor (T) antigen binds to sequences at the origin of viral DNA replication is not known. We (26, 27) and others (2, 33) have shown that the region of T antigen which is responsible for binding to DNA maps from residues 140 to about 260. This region does not appear to contain any previously characterized DNA-binding motif.

Papovavirus T antigens recognize the pentanucleotide sequence GAGGC at the origin of virus DNA replication (9–12, 22, 23, 34, 35). This sequence is present at T-antigen-binding sites I and II in SV40 DNA (9, 12, 34, 35). Sites I and II also contain A/T-rich tracts which induce DNA bending (7, 24). DNA bending is important for efficient binding of T antigen, especially at site I (24). In addition, site II, which constitutes the minimal core replication origin (7), contains an imperfect palindrome where DNA melting first occurs (4). Recently, Parsons et al. (21) have shown that T antigen can melt the imperfect palindrome in the absence of other origin sequences, suggesting that the protein interacts with it as well.

T antigen has a helicase activity (30) which may function at replication forks in unwinding parental strands. Recently (37), we showed that the helicase domain on T antigen (residues 131 to 616) extends from the beginning of the DNA-binding domain to the end of the ATPase domain. Our results (37) and those of others (3) strongly suggested that the DNA-binding domain of T antigen which is responsible for specific binding to the viral origin also binds nonspecifically to double-stranded and single-stranded DNA.

We have also recently demonstrated (29) that four major sequence elements of SV40 T antigen coordinate its specific and nonspecific DNA binding. These elements mapped to residues 152 to 155, 182 to 187, 203 to 207, and 215 to 219. The first three elements (A1, B1, and B2) were shown to be important for sequence-specific binding to sites I and II on the DNA, whereas the fourth element (B3) was found to be important in binding to site II only. On the basis of these observations, it was hypothesized that elements A1, B1, and B2 are required for binding to the GAGGC pentanucleotides and that element B3 binds to some other sequences in site II, perhaps to the imperfect palindrome. Element B1 was shown to be primarily involved in nonspecific binding to DNA.

For this report, we examined the first three elements in more detail by generating conservative mutations in and around each element. By testing the effects of these mutations on virus replication, DNA binding, and helicase activity, we determined the amino acid residues that are important for each activity. We have distinguished residues that are important for origin-specific binding from those that are involved in nonspecific binding and helicase activity. The information that we have obtained will be useful in understanding T-antigen-DNA interactions.

MATERIALS AND METHODS

Plasmids. pBS-SV40 contains the entire SV40 genome inserted into the *Bam*HI site of Bluescript (Stratagene) (16). pSKAT contains the SV40 T antigen gene inserted between adenovirus type 5 map units 0 to 1.4 and the major late promoter of adenovirus type 2 (29). pSVO+ contains T-antigen sites I and II, pOS1 contains site I, and pSVOdl3 contains site II (32).

Mutagenesis. Mutations were generated in pBS-SV40 or pSKAT by annealing oligonucleotides with a single mis-

* Corresponding author.

	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
SV40-	Phe	Leu	Ser	His	Ala	Val	Phe	Ser	Asn	Arg	Thr	Leu	Ala	Cys	Phe
BKV-	*	*	*	Gln	*	*	*	*	*	*	*	*	*	*	*
JCV-	*	*	*	Gln	*	*	*	*	*	*	*	Val	*	Ser	*
LPV-	*	*	*	*	*	Ile	Tyr	*	*	Lys	*	Met	Asn	Ser	*
HaPV-	*	*	*	*	*	Ile	*	*	*	Lys	*	Gln	Asn	Ala	*
PyV-	Tyr	*	*	*	*	Ile	Tyr	*	*	Lys	*	Phe	Pro	Ala	*
	Phe	Leu	Ser	His	Ala	Val	Phe	Ser	Asn	Arg	Thr	Leu	Ala		Phe
Mutation	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		↓
	Tyr	Val	Thr	Asn	Gly	Leu	Tyr	Thr	Ser	Lys	Ser	Ile	Gly		Tyr
Mutant name	145FY	146LV	147ST	148HN	149AG	150VL	151FY	152ST	153NS	154RK	155TS	156LI	157AG		159FY
Virus Rep	+	+	-	-	-	+	-	-	-	S	-	+	+		-
% DNA Binding D			43	68	6		71	49	32		33				22
% DNA Binding SS			76	68	18		88	67	125		116				7
% DNA Binding I			21	79	17		77	48	2		2				23
% DNA Binding II			27	78	13		101	32	0		0				4
Helicase			+	+	-		+	+	+		+				-

FIG. 1. Sequence and mutational analysis of element A. The amino acid sequences of element A in the T antigens of SV40, BK virus (BKV), JC virus (JCV), B lymphotropic papovavirus (LPV), hamster papovavirus (HaPV), and polyomavirus (PyV) are shown. The mutations that we made in this region and the mutant names are listed below the sequences. The effect of each mutation on virus replication was tested by transfection of mutant SV40 DNA into BSC-1 cells and then scoring for plaques (S, small plaques). T antigens were made corresponding to each replication-negative mutant and assayed for DNA-binding and helicase activities. Several different DNA substrates were used in binding reactions. D, 400-base-pair fragment of plasmid DNA to test for nonspecific binding to double-stranded DNA; SS, helicase substrate; I and II, T-antigen-binding sites I and II, respectively. Binding is represented as a percentage of that of wild-type T antigen.

match to a uridine-containing single-stranded DNA template as previously described (15, 16). The oligonucleotide was extended with T4 DNA polymerase (New England Bio-Labs), and the resulting double-stranded DNA was used to transform *Escherichia coli* BMH 71-18 (International Biotechnologies, Inc.). Single-stranded DNA was sequenced by the dideoxy procedure (25).

Virus replication assays. pBS-SV40 harboring a mutation in the T-antigen gene was cleaved with *Bam*HI to release the mutant genomic DNA. The DNA was ligated at low DNA concentrations to favor the formation of circular DNA and then transfected into monkey (BSC-1) cells by the DEAE-dextran procedure (18) as previously described (16). Plaques were counted 10 to 30 days posttransfection, depending upon plaque size. Plates which did not have any plaques were incubated for a minimum of 30 days to make sure that small plaques did not appear.

Preparation of mutant T antigen. Adenovirus-transformed 293 cells were transfected with *Kpn*I-linearized pSKAT and *Xba*I fragment A of adenovirus type 5 dl309 (13) (map units 4 to 100) by using a $CaPO_4$ precipitation technique (6). The presence of the adenovirus fragment stimulates T-antigen production by about twofold. Approximately 20 μg of pSKAT and 1.4 μg of *Xba*I fragment A were used per T75 flask of 293 cells. At 68 to 72 h, the cells were lysed and T antigen was recovered from the lysate by immunoprecipitation (28) with PAb416 monoclonal antibody. A sample of the immunoprecipitated T antigen was analyzed by acrylamide gel electrophoresis and Coomassie blue staining to quantitate the amount of protein.

DNA-binding assays. Quantitative DNA-binding assays were performed by incubating a ^{32}P-labeled *Taq*I fragment of pSVO+, pOS1, or pSVOdl3 with bound, immunoprecipitated T antigen by using a modification of the method of McKay (19), as previously described (28). Equal amounts of T antigen (about 0.1 μg) were used in each assay, and binding was measured at protein excess (about 4 ng of DNA per reaction). Binding of T antigen to site I (*Taq*I fragment E of pOS1) and site II (*Taq*I fragment E of pSVOdl3) were performed in the presence of a 1,000-fold mass excess of unlabeled calf thymus DNA as nonspecific competitor (29). Binding of T antigen to plasmid DNA was performed with ^{32}P-labeled *Taq*I fragment D of pSVO+ in the absence of unlabeled competitor DNA. Binding of T antigen to a ^{32}P-labeled helicase substrate (mostly single-stranded DNA) was also performed in the absence of competitor. The substrate was produced by hybridization of an oligonucleotide primer (15-mer) to single-stranded M13mp19 DNA and extending it to 19 nucleotides in the presence of [α-^{32}P] dATP, unlabeled dCTP, and Klenow polymerase as described previously (30). DNA-binding activity of mutant T antigen is reported as a percentage of the activity of wild-type T antigen.

DNase protection footprinting assays. DNase protection assays were performed under DNA replication conditions as described by Deb and Tegtmeyer (8). Fragments of T antigen (mostly containing sequences 131 to 708) were produced by trypsinization of immunoprecipitated T antigen (20 μg/ml for 30 min at 0°C) as previously described (26). T-antigen fragments (0.5 to 2 μg) were incubated with about 2 ng of a *Hin*dIII-*Nco*I fragment of pSVO+ containing T-antigen-binding sites I and II and which was labeled at the *Hin*dIII site. After 1 h at 37°C, the DNA was nicked with DNase (0.25 U/ml) for 5 min at 23°C and purified by phenol-chloroform extractions and ethanol precipitation. The DNA was denatured in formamide sample buffer and applied to a 7.5% acrylamide sequencing gel. As sequence markers, the same labeled DNA fragment was applied to the gel after treatment by the "G-only" reaction of Maxam and Gilbert (17).

Helicase assays. Helicase assays were performed by incubating solubilized tryptic fragments of T antigen (about 0.5 to 1 μg) with ^{32}P-labeled helicase substrate as described previously (30, 37). Labeled primer released from the helicase

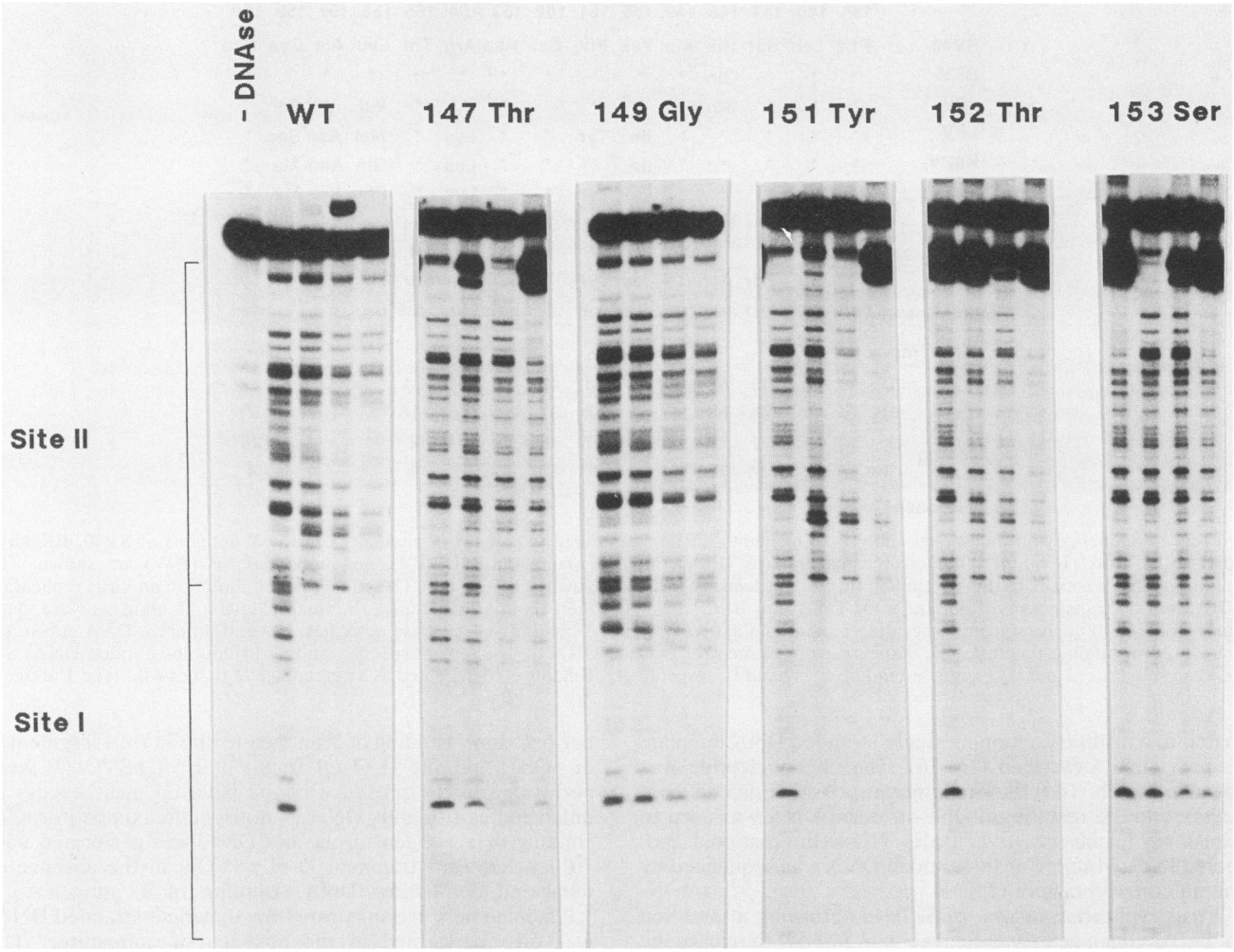

FIG. 2. DNase protection footprints of mutant T antigens from element A. Several representative mutant T antigens with changes in element A were purified by immunoprecipitation. Soluble fragments were generated by mild proteolysis and incubated with an end-labeled DNA fragment containing sites I and II. The DNA was nicked with a small amount of DNase I, denatured with formamide, and applied to a sequencing gel. One sample received no T antigen and no DNase (−DNAse). The four lanes under each T antigen correspond to samples which contained, from left to right, 0, 10, 20, and 40 μl of T antigen (corresponding to approximately 0 to 2 μg). Regions of the gel corresponding to DNA fragments that terminate in site I or II were determined by a "G only" sequencing reaction of the same end-labeled DNA and are shown on the left. The division between sites I and II is at nucleotide 5212. WT, Wild type.

substrate was detected by gel electrophoresis and autoradiography as described previously (30).

RESULTS

Mutational analysis of element A. As described in our previous report (29), we generated single-point substitution mutations in the DNA binding domain of SV40 T antigen. The mutations were chosen so as to make the mildest possible change (threonine for serine, lysine for arginine, leucine for isoleucine, etc.) in order to minimize effects on overall structure and stability. Our rationale was that mild mutations would most likely lead to the identification of amino acid residues which make contact with DNA.

Our earlier work (29) implicated three regions (called elements) within the domain in the binding to site I and four elements in the binding to site II at the SV40 replication origin. These conclusions were based on the effects of mutations generated at regular intervals within the domain. Mutations were concentrated in or around these elements, but not every site was mutated. In the present work, we generated mutations at nearly every position in and around elements A, B1, and B2. These three elements contain the sequences which are essential for the binding of T antigen to the recognition pentanucleotides at the origin.

Figure 1 shows the sequence of element A in SV40 T antigen and the corresponding sequences in the T antigens of five other papovaviruses. This region is fairly well conserved, especially between residues 145 and 155. Mutations were generated at every site, with the exception of position 158 (Fig. 1). The effect of each mutation on virus replication was tested by transfecting virus DNA into monkey cells and scoring for plaques. Mutants which did not replicate had changes at positions 147 to 149, 151 to 153, 155, and 159. All others replicated like wild-type virus with the exception of the mutant with a change at residue 154, which gave rise to small plaques. On the basis of these results, our previously defined element A1 (residues 152 to 155) (29) appears to

represent a portion of the larger element A (residues 147 to 159).

T antigens corresponding to replication-negative mutants were prepared by immunoprecipitation from transfected 293 cells and assayed for binding to several different DNA substrates and for helicase activity. Although DNA-binding activity was carefully quantitated, helicase activity was at times difficult to quantitate, and we therefore looked for the presence (+ in Fig. 1, 3, and 5) or absence (− in Fig. 1, 3, and 5) of a detectable displaced primer band (30). The phenotypes of the mutant proteins with changes at residues 149 and 159 were similar in that nonspecific binding to double-stranded (D) and single-stranded (SS, helicase substrate) DNA was seriously affected and in that both mutants lacked helicase activity (Fig. 1). Likewise, the two mutants with changes at residues 153 and 155 were similar in that they both had very low levels of specific binding to either site I or II on the DNA, whereas nonspecific binding and helicase activity were either unaffected or marginally affected. Mutations at position 147 and 152 had more moderate effects on specific DNA binding, but all other activities were similarly unaffected. Finally, two mutations (at residues 148 and 151) in this region had a minimal or no effect on DNA binding and helicase activity while having a major effect (at least a difference of 10^4 in titer) on virus replication.

The effects of mutations in this region on binding to sites I and II on the DNA was verified independently by performing DNase protection footprinting assays (Fig. 2). Unlike the DNA-binding assays done for Fig. 1, footprinting assays were performed under DNA replication conditions in the presence of ATP in order to maximize binding to site II (5, 8). DNase protection footprinting assays of some of the mutants with changes in element A are shown in Fig. 2. The results are consistent with those of the quantitative assay used in Fig. 1. The T antigen of mutants 153NS (153 Ser, Fig. 2) and 155TS (not shown) protected neither site I nor II, in agreement with the results shown in Fig. 1. Similarly, T antigens of mutants 147ST (147 Thr) and 149AG (149 Gly) protected both sites weakly while the T antigen of mutant 152ST (152 Thr) protected both sites somewhat better (compare with the numbers in Fig. 1). Finally, the T antigen of mutant 151FY (151 Tyr) and 148HN (not shown) protected both sites like wild-type T antigen.

These results implicate residues Ser-147, Ser-152, Asn-153, and Thr-155 in sequence-specific binding to both sites I and II at the origin. Of these, residues 153 and 155 appeared to be the most important, on the basis of the large effects of mutations at these two sites. On the other hand, the data implicate residues Ala-149 and Phe-159 in nonspecific binding to double- and single-stranded DNA and in helicase activity. Notice that all of these six amino acid residues are perfectly conserved in all papovavirus T antigens (Fig. 1).

Mutational analysis of element B1. We undertook a similar mutational analysis of element B1, which was shown in our previous work (29) to be important in nonspecific binding. On the basis of amino acid sequence homology with the closely related BK virus and JC virus T antigens, the limits of this region were thought to be 182 and 187. Our present work, however, suggest that this element spans only residues 185 to 187 because the mutants with changes at positions 182 to 184 replicated (Fig. 3). As described previously (29), mutations at residue 185 (185ST) and 187 (187HP and 187HR) had major effects on nonspecific (and origin-specific) binding to double-stranded DNA. Here we demonstrate that nonspecific binding to single-stranded DNA and helicase activity are also seriously affected (Fig. 3). Therefore, mutations at these two sites lead to a phenotype similar to those of some mutations in element A (at residues 149 and 159). Unlike the major changes in activity associated with mutations at residues 185 and 187, a mutation at position 186 had only slight effects on DNA-binding activity in the quantitative assay and no effect on helicase activity. DNase protection experiments (Fig. 4) showed that it had a more severe effect on DNA binding under replication conditions. Footprinting assays with mutant T antigen from 185ST (Fig. 4) confirmed its loss of sequence-specific DNA-binding activity (T antigen from the 187HP and 187HR mutants gave similar results; not shown).

		182	183	184	185	186	187		188	189
SV40-		Thr	Phe	Ile	Ser	Arg	His		Asn	Ser
BKV-		*	*	*	*	*	*		Met	Cys
JCV-		*	*	*	*	*	*		Gly	Phe
LPV-		Ser	*	Ser	Tyr	Gln	Asp		Glu	Glu
HaPV-		Ile	*	*	Met	Lys	Gln		Thr	Glu
PyV-		Glu	*	Lys	Cys	Leu	Val		His	Tyr
		Thr	Phe	Ile	Ser	Arg	His		Asn	Ser
Mutation		↓	↓	↓	↓	↓	↓	↘	↓	↓
		Ser	Tyr	Leu	Thr	Lys	Pro	Arg	Ser	Thr
Mutant name		182TS	183FY	184IL	185ST	186RK	187HP	187HR	188NS	189ST
Virus Rep		+	S	+	-	-	-	-	+	+
% DNA Binding	D				3	44	4	10		
	SS				0	55	16	6		
	I				18	59	1	4		
	II				12	68	3	8		
Helicase					-	+	-	-		

FIG. 3. Sequence and mutational analysis of element B1. T-antigen sequences corresponding to element B1 are shown and compared for several papovaviruses. T antigens were assayed for DNA-binding and helicase activities as shown in Fig. 1. Abbreviations are defined in the legend to Fig. 1.

Mutational analysis of element B2. Element B2 spans residues 203 to 207 and is highly conserved among papovavirus T antigens (Fig. 5). There is a change only at residue 206 in the hamster papovavirus T antigen. This position also appears to be the least important in the element on the basis of the fact that a mutation (Ser to Thr) there had no effect on virus replication (Fig. 5). Binding nonspecifically to double-stranded and single-stranded DNAs was most seriously affected in the mutant with a substitution at residue 203. The corresponding mutant T antigen was also deficient in helicase activity, placing it in the same class as two mutants (149AG and 159FY) in element A and three (185ST, 187HP, and 187HR) in element B1. Thus five sites have so far been implicated in nonspecific double-stranded and single-stranded DNA binding and in helicase activity. A mutation at residue 204 seriously affected specific binding to sites I and II with only a slight effect on nonspecific binding and a minimal effect on helicase activity, in agreement with our previous study (29). Changes at positions 205 and 207 affected specific DNA binding to sites I and II much less so (Fig. 5). This was confirmed by DNase protection assays (Fig. 4 and data not shown). Therefore, in region B2, Arg-204 appears to be the most important residue for sequence-specific binding to sites I and II.

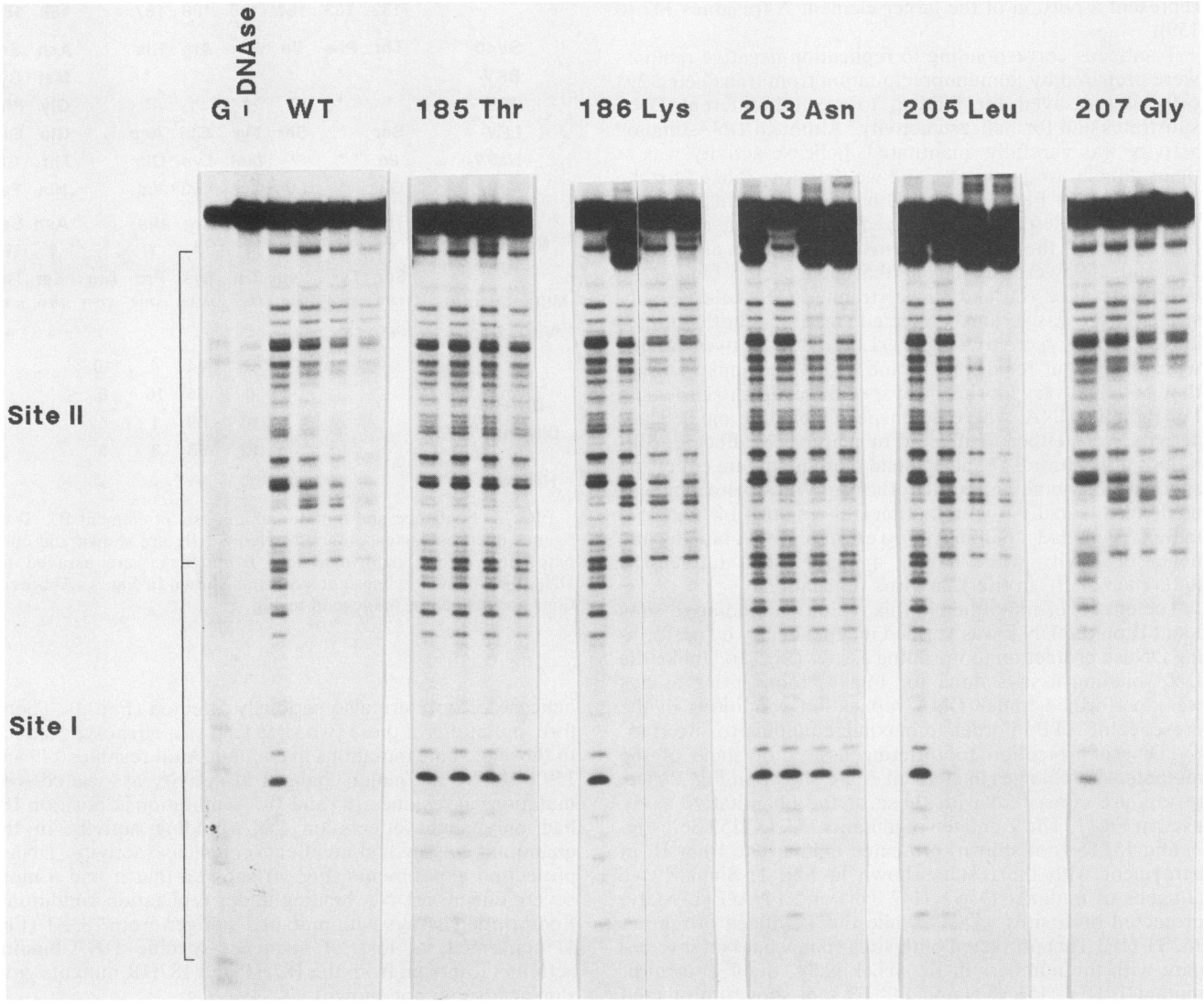

FIG. 4. DNase protection footprints of T antigen mutants from elements B1 and B2. Representative T antigens with changes in elements B1 or B2 were prepared and subjected to footprinting analysis as shown in Fig. 2. The lane marked "G" contains DNA reacted under "G only" sequencing conditions (17). Abbreviations are defined in the legend to Fig. 2.

DISCUSSION

Three regions of T antigen important for nonspecific and sequence-specific binding to DNA are shown in Fig. 6. Elements A and B2 consist of amino acid residues that are necessary for nonspecific or specific binding activity, while element B1 appears to contain residues involved only in nonspecific binding. Although other regions in the DNA-binding domain are important for overall structure and function, these three elements form the core of origin recognition.

On the basis of the magnitude of the effects of mutations at Asn-153, Thr-155, and Arg-204, these three residues seem to be the most important in the proper recognition of origin sequences (Fig. 6). Mutant T antigens with changes at these positions bound very poorly to both sites but were still capable of binding nonspecifically to double- and single-stranded DNAs and had helicase activity. Since both sites I and II contain the GAGGC recognition pentanucleotides, it seems likely that these three amino acids are involved in pentanucleotide binding. Although there are several possible ways in which they could mediate specific binding, one that we favor is that the residues are involved in direct or indirect contact with bases on the DNA. All three amino acids have the potential to make direct or solvent-mediated contacts with nucleotides, as determined by crystallographic studies of various procaryotic DNA-bindng proteins (1, 14, 20, 31). Threonine and several other residues (Ser, Pro, and Phe) have, as well, been shown to make van der Waals interactions with nucleotides in DNA (1).

In addition to the three residues listed above, several others are probably involved in binding to the pentanucleotide sequences, although they may be less likely to participate directly in the binding reaction. These are Ser-147 and Ser-152 in element A, and Val-205 and Ala-207 in element B2 (Fig. 6). Mutations at these sites affect sequence-specific binding to sites I and II, but the effects are not as dramatic as

		202	203	204	205	206	207	208
SV40-		Arg	His	Arg	Val	Ser	Ala	Ile
BKV-		*	*	*	*	*	*	*
JCV-		*	*	*	*	*	*	*
LPV-		Lys	*	*	*	*	*	Val
HaPV-		Lys	*	*	*	Ala	*	Val
PyV-		Lys	*	*	*	*	*	Val
		Arg	His	Arg	Val	Ser	Ala	Ile
Mutation		↓	↓	↓	↓	↓	↓	↓
		Lys	Asn	Lys	Leu	Thr	Gly	Leu
Mutant name		202RK	203HN	204RK	205VL	206ST	207AG	208IL
Virus Rep		+	-	-	-	+	-	+
% DNA Binding	D		3	29	49		25	
	SS		38	65	49		112	
	I		1	3	40		29	
	II		0	5	45		26	
Helicase			-	+	+		+	

FIG. 5. Sequence and mutational analysis of element B2. T antigen sequences corresponding to element B2 are shown and compared for several papovaviruses. T antigens were assayed for DNA-binding and helicase activities as shown in Fig. 1. Abbreviations are defined in the legend to Fig. 1.

with those involving residues 153, 155, and 204. Serine residues, being polar, can make hydrogen bonds with nucleotide bases (1, 14, 20, 31), so it is conceivable that residues 147 and 152 participate in binding. The effects of mutations at these two sites may not have been severe because the threonine that was introduced at each site substituted for the serine in making limited hydrogen bonding contacts. On the other hand, the hydrophobic amino acids at residues 205 and 207 are less likely to be directly involved in binding, although there is a report (20) that the peptide nitrogen of some hydrophobic residues (Ile and Ala) can form a water-mediated H bond with adenine or guanine bases. These sites may therefore be utilized in indirect contacts but perhaps are more likely required for the proper positioning of contact residues. It is also possible that these residues are important for nonspecific DNA binding since this activity was reduced in the mutants. However, in our previous study (29), both

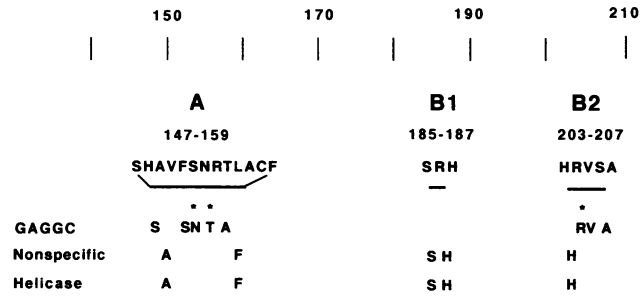

FIG. 6. Summary of important amino acid residues in elements A, B1, and B2. Residues implicated in binding to the recognition pentanucleotide GAGGC, in nonspecific binding to double- and single-stranded DNAs, and in helicase activity are highlighted. Single-letter codes are used to designate amino acids. The asterisks above N, T, and R signify that these three residues (Asn-153, Thr-155, and Arg-204) are thought to be the most important in specifying binding to the pentanucleotides.

mutant proteins bound DNA nonspecifically at close to wild-type levels, and in the present study, neither mutant protein had the very low nonspecific double-stranded DNA-binding activity characteristic of others (see below). Further, the mutation at residue 207 had no effect on single-stranded DNA binding.

Our results implicate five amino acids in nonspecific binding to DNA. Mutations at Ala-149, Phe-159, Ser-185, His-187, and His-203 all resulted in T antigens that were seriously affected in their ability to bind nonspecifically to double-stranded and single-stranded DNAs and in helicase activity. The first two residues implicated in nonspecific binding lie in element A, the next two lie in element B1, and the last one lies in B2. It seems reasonable to postulate that these residues are involved in the direct or indirect interaction with the sugar-phosphate backbone of DNA. Basic amino acids (like histidine) and polar amino acids (like serine) have the potential to make contact with the backbone, as shown by X-ray diffraction analysis of various protein-DNA cocrystals (1, 14, 20, 36). Less is known about the involvement of hydrophobic amino acids in backbone contacts. However, Jordan and Pabo (14) have shown that the peptide N of Ala-56 of the lambda repressor makes a H bond with a phosphate. Since His-203 is very close to the important Arg-204, this histidine could be involved in specific as well as nonspecific binding. Examples of this kind exist. In the lambda repressor (14), Gln-33 forms an H bond with a phosphate on the DNA and another with Gln-44, which itself makes two H bonds with an adenine.

The results described in this paper are generally consistent with those of our previous report (29) with the exception of those obtained with the 149AG mutant, which was reported earlier to bind DNA like wild-type T antigen. Our present findings have been confirmed several times with different mutant isolates, and we believe that the previous experiments were performed with a recombinant adenovirus containing a wild-type revertant of the T-antigen gene. In our present work, T antigen was produced by direct transfection of mutant DNA, not by infection with recombinant viruses, so there was a smaller chance of picking up a second mutation.

We note that each residue in elements A and B2 which is implicated in specific or nonspecific binding is completely conserved in all known papovaviruses (Fig. 1 and 5). This observation supports our contention that these sites are important. In element A, Arg-154 was mutated to a Lys and the mutant virus replicated but gave rise to small plaques (Fig. 1). At that position, the residue was Arg in the T antigens of SV40, BK virus, and JC virus but Lys in those of the other papovaviruses. Therefore, it is conceivable that a basic residue (either Arg or Lys) is necessary here. Mutation of this residue to a Ser abolishes virus replication (not shown). The same situation might apply at position 186, where Lys could partially substitute in DNA binding for the Arg at that site in SV40 T antigen. Unlike important residues in elements A and B2, those in element B1 (Ser-185 and His-187) appear to be conserved only in the closely related papovaviruses SV40, BK virus, and JC virus (Fig. 3). We believe that these two residues are important in nonspecific contacts with DNA and that equivalent amino acids should be involved in the other three papovavirus (lymphotropic papovavirus, hamster papovavirus, and polyomavirus) T antigens as well. Polyomavirus T antigen is actually one residue shorter between elements A and B1, so structurally, the residues corresponding to positions 185 and 187 of SV40 T antigen would be listed under positions 186 and 188 in the

polyomavirus sequence shown in Fig. 3. The histidine in the polyoma sequence under position 188 (actually residue 341) might therefore perform the same function as His-187 of SV40 T antigen.

We hypothesize that the polypeptide chains corresponding to elements A and B2 are positioned in the major groove of DNA and participate in the proper recognition of $\frac{5'\ GAGGC\ 3'}{3'\ CTCCG\ 5'}$ base pairs. It has been suggested that the specific DNA recognition domain of T antigen fits in the major groove of DNA (12). Since one complete turn of B-DNA is 10 base pairs, GAGGC pentanucleotides would occupy one-half turn or one major groove. According to our model, certain residues in elements A and B2 are accomodated in the major groove and are involved in pentanucleotide binding. Of the five base pairs, the third and fourth G · Cs and the last C · G are probably the most important (9). Dimethylsulfate protection experiments (9) showed that all guanine bases within pentanucleotides probably make contact with T antigen. At this time, we are not able to assign possible amino acid-base contact pairs, since Arg, Thr, and Asn are capable of H bonding several different bases, including guanine. Besides, one cannot exclude the possibilities that some contacts are water or cation mediated (20) and that specific interactions are due to van der Waals forces, not H bonds.

The correlation between nonspecific DNA binding and helicase activity (Fig. 6) indicates that all nonspecific contacts with DNA are required for helicase activity. This suggests that the helicase is dependent on the continuous interaction with DNA. This activity is, however, totally independent of specific DNA binding (Fig. 6), in agreement with previous observations (3).

In conclusion, we have identified the residues which are most likely involved in specific binding to pentanucleotide sequences at the origin and those involved in nonspecific binding and helicase activity within the DNA-binding domain (Fig. 6). Although we cannot exclude the possibility that other unidentified residues function directly in pentanucleotide binding, we believe that most of them are described in this study. How these amino acids are oriented relative to the DNA is a question of great interest which, however, can only be answered by X-ray crystallography of T-antigen-DNA cocrystals.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant CA 36118 from the National Cancer Institute.

LITERATURE CITED

1. **Aggarwal, A. K., D. W. Rodgers, M. Drottar, M. Ptashne, and S. C. Harrison.** 1988. Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. Science **242**:899–907.
2. **Arthur, A. K., A. Höss, 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.
3. **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.
4. **Borowiec, J., and J. Hurwitz.** 1988. Localized melting and structural changes in the SV40 origin of replication induced by T-antigen. EMBO J. **7**:3149–3158.
5. **Borowiec, J., 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. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**:2745–2752.
7. **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.
8. **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.
9. **DeLucia, A. L., B. A. Lewton, R. Tijan, and P. Tegtmeyer.** 1983. Topography of simian virus 40 A protein-DNA complexes: arrangement of pentanucleotide interaction sites at the origin of replication. J. Virol. **46**:143–150.
10. **Dilworth, S., A. Cowie, R. Kamen, and B. Griffin.** 1984. DNA-binding activity of polyoma virus large tumor antigen. Proc. Natl. Acad. Sci. USA **81**:1941–1945.
11. **Gottlieb, P., M. S. Nasoff, E. F. Fisher, A. M. Walsh, and M. H. Caruthers.** 1985. Binding studies of SV40 T-antigen to SV40 binding site II. Nucleic Acids Res. **13**:6621–6634.
12. **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.
13. **Jones, N., and T. Shenk.** 1979. Isolation of adenovirus type 5 host range deletion mutants defective for transformation of rat embryo cells. Cell **17**:683–689.
14. **Jordan, S. R., and C. O. Pabo.** 1988. Structure of the lambda complex at 2.5 Å resolution: details of the repressor-operator interactions. Science **242**:893–899.
15. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
16. **Loeber, G., R. Parsons, and P. Tegtmeyer.** 1989. The zinc finger region of simian virus 40 large T antigen. J. Virol. **63**:94–100.
17. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
18. **McCutchan, J. H., and J. S. Pagano.** 1968. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyldextran. J. Natl. Cancer Inst. **41**:351–357.
19. **McKay, R.** 1981. Binding of simian virus 40 T antigen-related protein to DNA. J. Mol. Biol. **145**:471–488.
20. **Otwinowski, Z., R. W. Schevitz, R.-G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi, and P. B. Sigler.** 1988. Crystal structure of trp repressor/operator complex at atomic resolution. Nature (London) **335**:321–329.
21. **Parsons, R., M. E. Anderson, and P. Tegtmeyer.** 1990. Three domains in the simian virus 40 core origin orchestrate the binding, melting, and DNA helicase activities of T antigen. J. Virol. **64**:509–518.
22. **Pomerantz, B. J., and J. A. Hassell.** 1984. Polyomavirus and simian virus 40 large T antigens bind to common DNA sequences. J. Virol. **49**:925–937.
23. **Ryder, K., A. L. DeLucia, and P. Tegtmeyer.** 1983. Binding of SV40 A protein to the BK virus origin of DNA replication. Virology **129**:239–245.
24. **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.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
26. **Simmons, D. T.** 1986. DNA-binding region of the simian virus 40 tumor antigen. J. Virol. **57**:776–785.
27. **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.
28. **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.
29. **Simmons, D. T., G. Loeber, and P. Tegtmeyer.** 1990. Four major sequence elements of simian virus 40 large T antigen coordinate

its specific and nonspecific DNA binding. J. Virol. **64**:1973–1983.

30. **Stahl, H., P. Droege, and R. Knippers.** 1986. DNA helicase activity of SV40 large tumor antigen. EMBO J. **5**:1939–1944.
31. **Steitz, T. A., and C. M. Joyce.** 1987. Exploring DNA polymerase I of *E. coli* using genetics and X-ray crystallography, p. 227–235. *In* D. L. Oxender and C. F. Fox (ed.), Protein engineering. Alan R. Liss, Inc., New York.
32. **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.
33. **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.
34. **Tjian, R.** 1978. Protein-DNA interactions at the origin of simian virus 40 DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:655–662.
35. **Tjian, R.** 1978. The binding site of SV40 DNA for a T antigen-related protein. Cell **13**:165–179.
36. **Wolberger, C., Y. Dong, M. Ptashne, and S. C. Harrison.** 1988. Structure of a phage 434 cro/DNA complex. Nature (London) **335**:789–795.
37. **Wun-Kim, K., and D. T. Simmons.** 1990. Mapping of helicase and helicase substrate binding domains on simian virus 40 large T antigen. J. Virol. **64**:2014–2020.