JOURNAL OF VIROLOGY, Dec. 1990, p. 6291–6296
0022-538X/90/126291-06$02.00/0

Vol. 64, No. 12

The Finger Domain of Simian Virus 40 Large T Antigen Controls DNA-Binding Specificity

ADOLF HÖSS, ISMAIL F. MOAREFI, ELLEN FANNING, AND AVRIL K. ARTHUR*
Institut für Biochemie, Karlstrasse 23, 8000 Munich 2, Federal Republic of Germany

Received 24 April 1990/Accepted 21 August 1990

The specificity and regulation of protein-DNA interactions play a crucial role in all aspects of communication between genotype and phenotype in a cell. The large T antigen of simian virus 40 binds to identical, yet quite differently arranged, pentanucleotide motifs in the simian virus 40 control region, sites I and II. Wild-type T antigen preferentially binds site I. We demonstrate that a bacterial peptide encoding residues 1 to 259 (T260) includes the essential amino acids required for binding to both DNA sites but predominantly binds site II. However, a longer peptide (residues 1 to 369; T370) binds almost exclusively to site I. Thus, the addition of amino acids 260 to 369 to the T260 peptide results in the loss of site II binding. This region includes a single putative metal-binding region, and mutation of T370 at either conserved cysteine of the finger results in equal but inefficient binding to both sites. While no metal binding has been shown to be directly associated with this sequence, these results suggest a novel, perhaps structural, function for such a finger motif, since this domain of T antigen appears to play a crucial role in modulating the DNA-binding behavior of T-antigen peptides.

T-antigen origin DNA-binding activity is highly controlled since only a fraction of the protein present in lytically infected cells has specific DNA-binding activity (24). In addition, differential binding to either site I or site II is influenced by numerous parameters, including phosphorylation, oligomerization, ATP, and Mg^{2+} (reviewed in references 8 and 27), but what determines how T antigen binds with different activities to either site is still not understood. It is known, however, that T-antigen binding to site II, which contains two pairs of pentanucleotide in inverted orientation, is essential for the initiation of simian virus 40 (SV40) DNA replication, while the minimal site I, involved primarily in autoregulation of early transcription, contains two of the same pentanucleotides as direct repeats (Fig. 1A) (reviewed in references 6 and 27). Thus, the two sites are not only involved in different functions but are presented in a very different DNA conformational context: site II with each pair of inverted repeats appropriately spaced to lie on one side of a normal B-DNA helix, and the site I direct repeat motif present as bent DNA due to the A+T-rich spacer (21) (Fig. 1A).

We and others have previously located the DNA-binding domain of T antigen for both sites between amino acids 131 and 259 (1, 18, 27, and references therein). While it is clear that these sequences are both necessary and sufficient for specific binding to both sites I and II, data obtained with longer peptides (1, 18) indicated that additional regions of T antigen play a role in modifying the DNA-binding activity. This prompted us to investigate further the effect of sequences outside of the DNA-binding domain on the DNA-binding profile with the aim of characterizing features essential for efficient binding to either site I or site II in the SV40 control region.

Site I and site II DNA binding. We analyzed the DNA-binding profile of full-length T antigen and several truncated T-antigen peptides (depicted schematically in Fig. 1B) expressed in *Escherichia coli*, either as crude extracts or after immunopurification to near homogeneity (Fig. 2) (1). Expression vectors, bacterial growth conditions, preparation of *E. coli* crude extracts, and purification of peptides expressed in *E. coli* were essentially as previously described (1, 13). From 500 ml of cells, approximately 2 mg of greater than 90% pure T260 or 0.1 mg of pure T370 was obtained. All purified proteins were analyzed by immunoblotting with the appropriate monoclonal antibodies to confirm the expected T-antigen epitopes (data not shown). Cell lysate supernatants from induced *E. coli* cultures (6 ml of T1, T302, T305; 2 ml of all others) or 250 ng of purified baculovirus T antigen or 200 ng of purified bacterial mutant peptides were incubated with equimolar amounts (50 fmol) of restricted, end-labeled pONwt and p1097 or pSVwt DNA (Fig. 1A) in binding buffer (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid] [pH 7.8, 50 mM NaCl, 2 mM EDTA, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 8% glycerol, 0.2 mg of glycogen per ml, 1 mg of bovine serum albumin per ml, 1 μg of salmon sperm DNA per ml). After 1 h at 4°C, bound complexes were immunoprecipitated with monoclonal antibody Pab419 (12) or, for T131-260, Pab220 (gift from S. Mole and D. Lane). The bound DNA was separated on 1.5% agarose gels and visualized by autoradiography as described previously (1, 26).

In all cases tested, the material from the crude extracts bound the DNA fragments in a manner indistinguishable from that of the purified material (Fig. 3, compare panels A and C). The control protein in these experiments was T antigen purified from insect cells (recombinant baculovirus), the expression and purification of which have been described elsewhere (13a). Baculovirus T antigen bound both site I- and site II-containing fragments, but preferably site I, in a manner similar to that of T antigen from mammalian cells (Fig. 3A, lane 4) as reported previously (15, 20). Full-length T antigen expressed in *E. coli*, however, only detectably bound site I (Fig. 3A, lane 3) as previously described (19). The failure to bind site II has been ascribed to the underphosphorylated nature of this protein (17, 19). The bacterial

* Corresponding author.

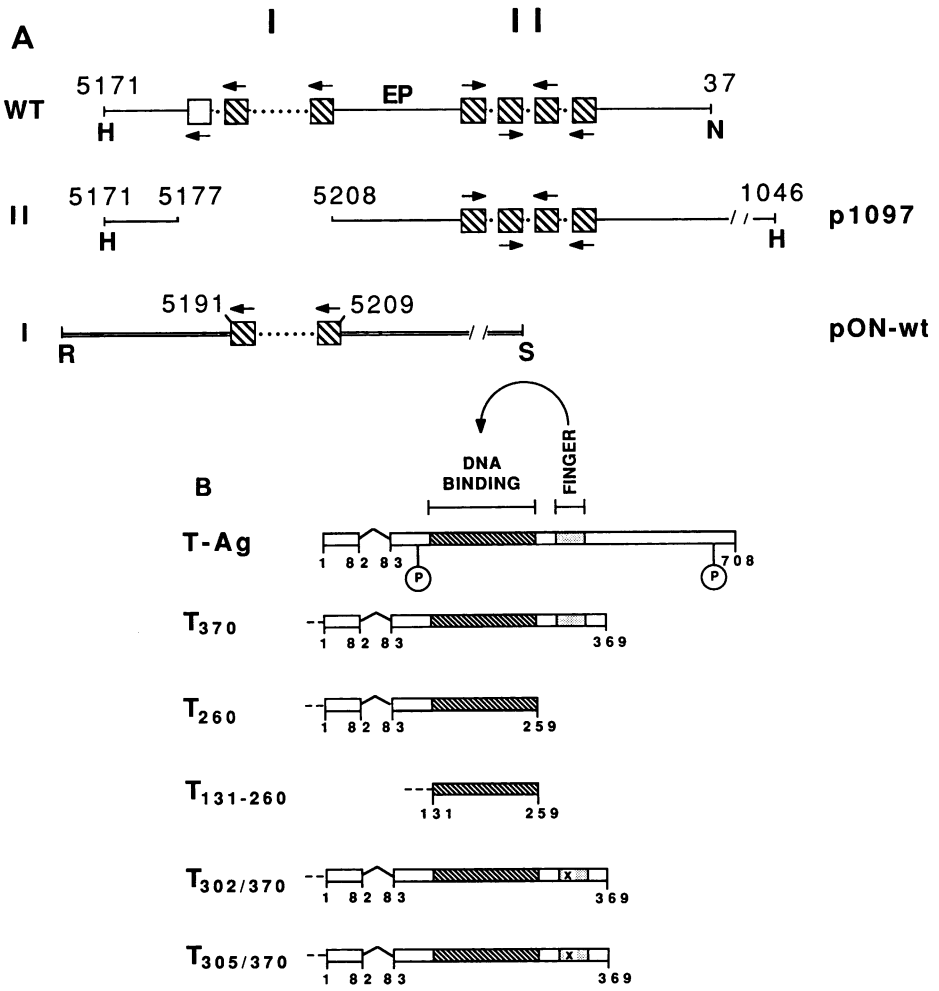

FIG. 1. (A) Top (WT): Diagram of the arrangement of the perfect pentanucleotides GAGGC (shaded boxes) and one imperfect pentanucleotide (empty box) found in site I and site II of the SV40 control region, their direction and side of the DNA helix (indicated by the arrows), and the early palindrome (EP) located between them. Site II contains two pairs of perfect inverted pentanucleotides separated by 7 base pairs, while site I contains two perfect pentanucleotides in direct repeat separated by a 7-base-pair A+T-rich spacer (dotted region) which confers a bend in the DNA (21). The sequence between the *Hin*dIII (H; nucleotide 5171) and *Nco*I (N; nucleotide 37) sites was the substrate for DNase I footprinting of the wild-type origin (see Fig. 4). Below: Substrates used in the McKay DNA-binding assays (see Fig. 3A, C, and D). The site II substrate (middle) was a *Hin*dIII (H; nucleotides 5171 to 1046) digest of p1097 which carries a deletion of 31 base pairs encompassing all of site I (7). The site I substrate (bottom) was an *Eco*RI (R)-*Sal*I (S) fragment from pONwt (22) which contains a 19-base-pair oligonucleotide with the two shaded pentanucleotides plus spacer region from site I cloned into pAT153 (double lines), and which is bound with affinity equal to wild-type site I (22). The wild-type origin fragment (WT) for the McKay assays (see Fig. 3B) is the 1,117-base-pair *Hin*dIII fragment of pSVwt (pAT153 containing the entire SV40 genome) (26). (B) Diagram of the peptides described in this report with T-antigen residues numbered below. The peptides were cloned and expressed from the *lac* promoter of pUC9 as described previously (1). N-terminal fusion amino acids are indicated by dashes. Stop codons, novel restriction sites to remove the intron, and amino acid substitutions (Cys-302 or -305 to Ser [X]) were introduced by oligonucleotide site-directed mutagenesis as described previously (1, 26). The DNA-binding domain and putative finger region are shown above, and the clustered phosphorylated residues (P) are shown below (23). T-Ag, T antigen.

peptide T370 (residues 1 to 369) bound the mixture of fragments in a manner identical to that of the full-length bacterial protein (Fig. 3A, compare lanes 2 and 3). Contrary to this, T131-260 (residues 131 to 259) bound both sites but preferred site II (Fig. 3A, lane 5). Similarly, a peptide including the first 259 residues of T antigen, T260, bound preferentially to site II, although site I binding was also detected with higher amounts of protein (Fig. 2A or C, lane 1).

This result is in direct contrast to previously reported data (17) which showed that an almost identical peptide bound well to site I but very poorly to site II unless phosphorylated in vitro by *cdc-2* kinase. This difference may stem from variations between the assay conditions, since we observed preferential site II binding with that peptide (gift from D. McVey). Therefore, to confirm the site II DNA-binding specificity of the peptide T260, we performed alternative assays, including DNase I footprint analyses of the wild-type origin region. Purified baculovirus wild-type T antigen, used as a control, protected site I with low amounts of protein, while with greater amounts, full protection extended through the early palindrome and site II and sometimes into the A+T-rich region (Fig. 4A and B). In contrast, T260 protected site II with the lowest amount of protein and with

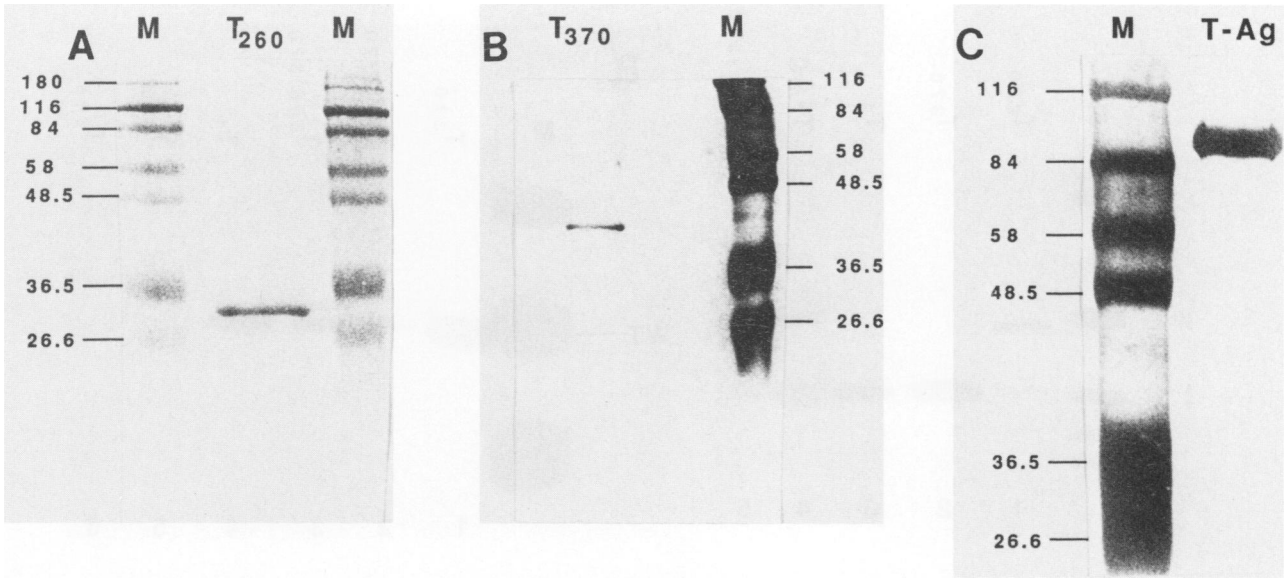

FIG. 2. Polyacrylamide gel electrophoretic analysis of T antigen and peptides purified by immunoaffinity chromatography. (A) *E. coli*-expressed peptide T260, Coomassie blue-stained gel. (B) *E. coli*-expressed peptide T370, silver-stained gel. (C) Baculovirus-expressed full-length T antigen, silver-stained gel. Lanes M, Prestained protein size markers (Sigma). Sizes are shown by numbers on sides of gels (in kilodaltons).

increasing amounts partially protected site I, but failed to protect the early palindrome between (Fig. 4B). While this may be because T260 is a much smaller protein than full-length T antigen, its failure to protect the early palindrome is not due to an inability of T260 to form oligomers (A. Höss and A. K. Arthur, unpublished data). Filter binding experiments with isolated site I or site II DNA restriction fragments showed that T260 bound site II with a greater affinity than site I (Höss and Arthur, unpublished data). These results confirm the specific and preferential binding of T260 to site II.

Influence of finger domain on DNA binding. The only other mutations so far described which modulate the DNA-binding activity of T antigen, particularly site II binding, are conservative amino acid changes of potentially phosphorylated serine or threonine residues clustered at the N and C termini of the protein (8, 26) (Fig. 1B). However, by expressing our mutants in *E. coli*, we are essentially analyzing unphosphorylated peptides (A. Arthur, unpublished data) as demonstrated by their being excellent substrates for various cellular kinases in vitro (13a). In addition, the peptide T131-T260, which contains none of the potentially phosphorylated residues (Fig. 1B), bound site II efficiently, in a manner equivalent to that of the T260 peptide, which does contain the N-terminal cluster of phosphorylation sites. However, the peptides T260 and T370 bear identical N termini (Fig. 1B) but bind sites I and II quite differently, implicating the sequence between residues 259 and 369 as being responsible for the results reported here.

Located between residues 260 and 370 is a possible metal-binding sequence (14, 16), which is highly conserved among six related papovavirus T antigens (J. M. Pipas, personal communication). Replacement of either cysteine residue 302 or 305, which would be involved in forming a finger structure, by serine reduced specific origin DNA-binding activity of full-length T antigen to levels which are undetectable by this assay (Fig. 3B, lanes 4 to 6) (1). These mutations have also been shown to greatly reduce T-antigen-mediated SV40 DNA replication in monkey cells and transformation in rat cells (16). When these mutations were introduced into the truncated T370 peptide, specific DNA binding was retained (T302/370, T305/370; Fig. 3B, lanes 2 and 3). However, it is clear that the overall binding capacity of these mutants was greatly reduced, since they bound only 5 to 10% of the origin fragment that was bound by the T370 peptide or full-length T antigen (determined by densitometry of the autoradiogram in Fig. 3B; data not shown), despite using appropriate volumes of each extract such that equivalent amounts of T antigen and peptides were present in all assays (as determined by immunoblot; data not shown).

When binding to the individual sites I and II was tested with equivalent amounts of purified mutant T370 proteins, binding to site I was found to be considerably reduced compared with wild-type T370, while binding to site II was increased (Fig. 3D). These mutations, therefore, neither confer the efficient site II binding which results from the deletion of the finger region, as demonstrated by the T260 mutant, since detectable site II binding required a longer exposure of the autoradiograph (compare Fig. 3C and D), nor do they retain the efficient site I binding of their progenitor, T370. The DNA-binding profile of the mutants T302/370 and T305/370 is thus intermediate between that of T260 and T370 and may be due to the single amino acid substitution partially disrupting the finger, which could then lead to reduced DNA binding, with no strong preference for either site.

There is still no conclusive evidence that the finger sequence of T antigen specifically binds any metal ion. Obviously, vigorous testing of purified peptides, both biochemically and structurally, will be required to determine whether this sequence is indeed a metal-binding finger. Nevertheless, our data show that both cysteines 302 and 305 appear crucial for some structural feature of T antigen, since their deletion

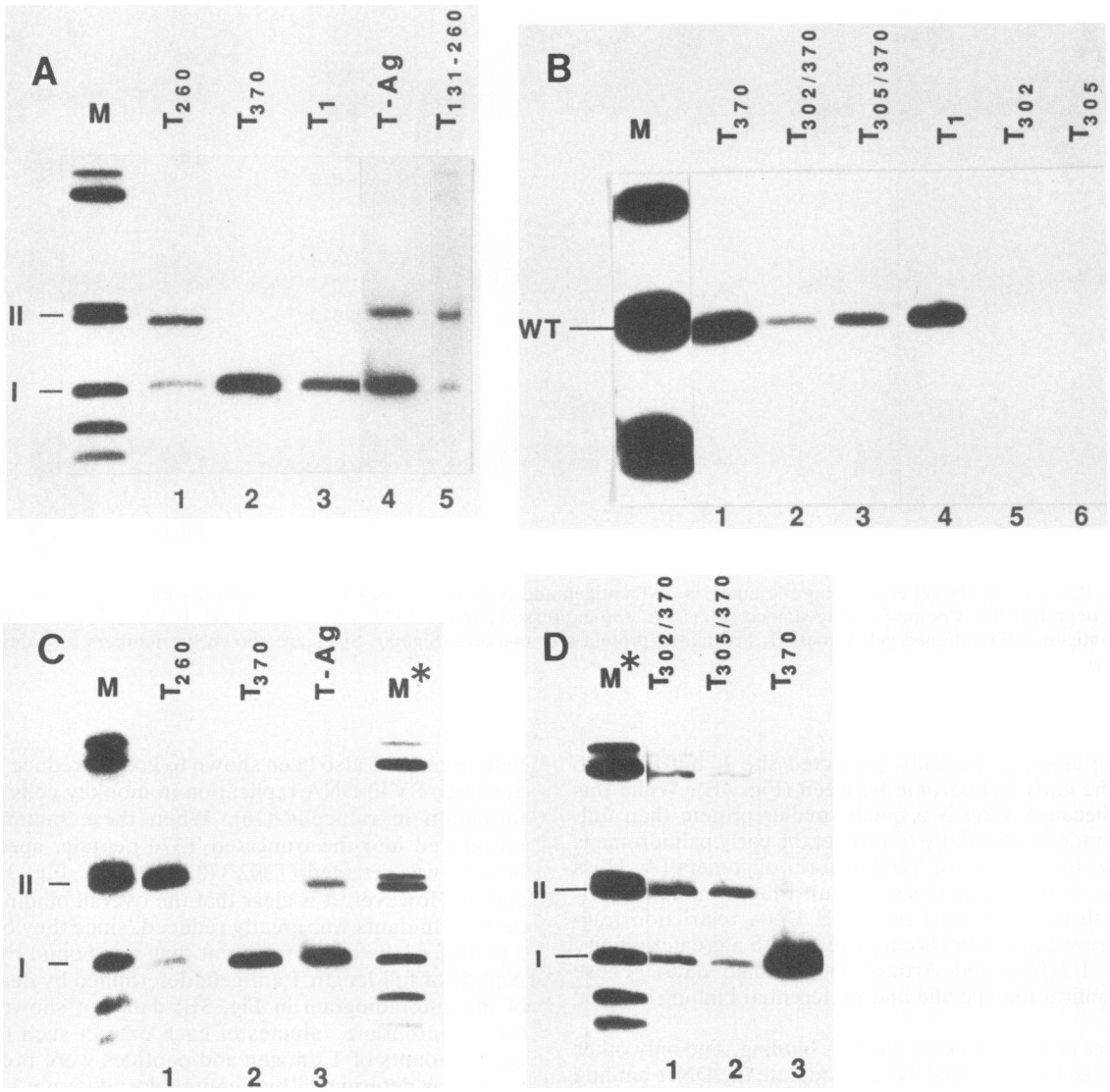

FIG. 3. (A and B) Specific DNA binding of *E. coli* extracts containing mutant peptides or full-length T antigen (T1). Immunoaffinity-purified recombinant baculovirus T antigen was tested as a control (T-Ag). Site I-, site II-, or wild-type origin-containing fragments are indicated. (B) A long exposure of the autoradiogram (wild-type origin binding) is shown so that binding in lanes 2 and 3 (T302/370, T305/370) is visible. No specific bands were detected in lanes 5 or 6 (full-length bacterial T antigen containing Cys-Ser substitutions at position 302 or 305; see text), regardless of the exposure time. Equivalent amounts of protein expressed by all constructs were used, as determined by immunoblot (data not shown). (C and D) DNA binding of immunoaffinity-purified peptides from *E. coli* or the wild-type baculovirus T antigen (T-Ag). (C) Lane M, 20% of the input DNA used in the binding reactions, 6-h film exposure time; lane M*, a shorter exposure of the marker lane to show the individual restriction fragments. (D) Overnight exposure of lanes 1 to 3; lane M*, a shorter exposure of the marker lane.

or substitution has such a profound effect on the DNA-binding activity of the protein.

Metal-binding finger motifs have been implicated, and in some cases proved to be directly involved in sequence-specific protein-DNA interactions of transcription factors (reviewed in references 2, 9, and 14). However, it is also becoming clear that they may often perform other roles related to protein conformation, stability, and protein-protein interactions (9, 28), for example, being essential for the activity but not the sequence specificity of GAL4 DNA binding (5) or, in the case of the second finger of the glucocorticoid-estrogen receptors, the protein-protein interactions between subunits (2, 11). Other functions include non-sequence-specific interactions with DNA, as, for example, in the gene 32 protein of bacteriophage T7, which binds single-stranded DNA (10), or adenovirus E1A (4); double-stranded RNA binding of a retroviral *gag* protein (25); or primase activity of the T7 gene 4 (3).

Our results suggest that the finger sequence of T antigen does not play a direct role in contacting the pentanucleotides of either site I or II, since the T260 peptide can bind both sites specifically. Instead, the wild-type finger sequence seems to modulate the DNA-binding specificity of the protein, either interfering with site II binding or favoring site I binding. While it has yet to be tested whether the altered DNA-binding preferences of T260 and T370 are due to protein-protein interactions or to a protein conformational change induced by the finger structure, the mutants de-

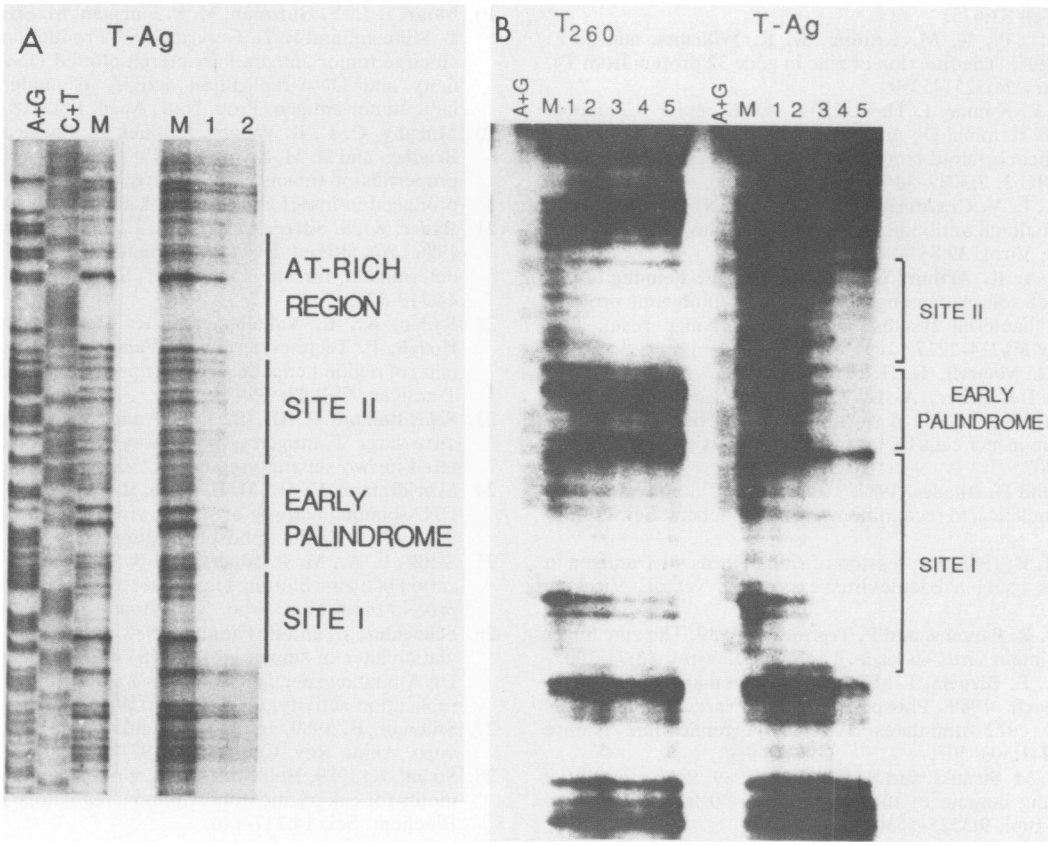

FIG. 4. DNase I protection analysis of a wild-type SV40 origin fragment (Fig. 1A, WT) with purified bacterial T260 or recombinant baculovirus wild-type T antigen. The single end-labeled (^{32}P) fragment in panel A was subjected to Maxam-Gilbert sequencing reactions (A+G and C+T) with Merck kit (catalog no. 6889) or digested with DNase I in the absence of protein (lane M) or after incubation with 500 ng (lane 1) or 10 mg (lane 2) of wild-type baculovirus T antigen. Reactions were incubated with 20 fmol of the wild-type origin fragment (144 base pairs) for 20 min at room temperature in 25 ml of binding buffer with 500 ng of sonicated salmon sperm DNA as the competitor. DNase I (Pharmacia) was added to 50 pg/ml in a 70-ml reaction mixture (final volume; 10 mM HEPES [pH 7.8], 10 mM $MgCl_2$, 5 mM $CaCl_2$) for 1 min at room temperature. The reaction was stopped by adding 10 mM EDTA, 0.3 M sodium acetate, and 50 mg of salmon sperm DNA per ml. The DNA was phenol extracted, ethanol precipitated, and analyzed on an 8% DNA-sequencing gel. (B) Incubation with increasing amounts of T260 (lane 1, 200 ng; lane 2, 500 ng; lane 3, 1,250 ng; lane 4, 2.5 μg; lane 5, 5 μg) or baculovirus T antigen (lane 1, 200 ng; lane 2, 500 ng; lane 3, 2.0 μg; lane 4, 5 μg; lane 5, 20 μg) protein before digestion with DNase I.

scribed here provide a relatively simple system to test directly which residues, modifications, or amino acid sequences influence the capacity of peptides to bind specifically to site I or II.

We thank Ronald Mertz, Doris Weigand, and Andrea Schweizer for oligonucleotide synthesis; Haralabos Zorbas for invaluable technical advice and interest; Stephan Glaser for critical reading of the manuscript; Silke Dehde and Birgit Posch for supplying media, antibodies, and other materials; and Irene Dornreiter for advice on the purification of baculovirus T antigen. We gratefully acknowledge Sara Mole and David Lane for kindly supplying monoclonal antibodies, Duncan McVey for supplying the 259.3 peptide, and all members of our group for helpful discussions.

A.H., I.F.M., and A.K.A. were supported by the Bundesministerium für Forschung und Technologie (Genzentrum), and additional funding for this work came from the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

LITERATURE CITED

1. **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 to within 129 amino acids. J. Virol. **62**:1999–2006.
2. **Berg, J. M.** 1989. DNA binding specificity of steroid receptors. Cell **57**:1065–1068.
3. **Bernstein, J. A., and C. C. Richardson.** 1988. A 7-kDa region of the bacteriophage T7 gene 4 protein is required for primase but not for helicase activity. Proc. Natl. Acad. Sci. USA **85**:396–400.
4. **Chatterjee, P. K., M. Bruner, S. J. Flint, and M. L. Harter.** 1988. DNA-binding properties of an adenovirus 289R EIA protein. EMBO J. **7**:835–841.
5. **Corton, J. C., and S. A. Johnston.** 1989. Altering DNA-binding specificity of GAL4 requires sequences adjacent to the zinc finger. Nature (London) **340**:724–727.
6. **DePamphilis, M. L., and M. K. Bradley.** 1986. Replication of SV40 and polyoma virus chromosomes, p. 99–246. *In* N. P. Salzman (ed.), The papovaviridae, vol. 1. Plenum Publishing Corp., New York.
7. **DiMaio, D., and D. Nathans.** 1982. Regulatory mutants of simian virus 40. Effect of mutations at a T-antigen binding site on DNA replication and expression of viral genes. J. Mol. Biol. **156**:531–548.
8. **Fanning, E., J. Schneider, A. Arthur, A. Höss, I. Moarefi, and S. Modrow.** 1989. Structure and function of SV40 large T antigen communication between functional domains. Curr. Top. Microbiol. Immunol. **144**:9–19.
9. **Frankel, A. D., and C. O. Pabo.** 1988. Fingering too many

proteins. Cell **53**:675.
10. **Giedoroc, D. P., K. M. Keating, K. R. Williams, and J. E. Coleman.** 1987. The function of zinc in gene 32 protein from T4. Biochemistry **26**:5251–5259.
11. **Green, S., V. Kumar, I. Theulaz, W. Wahli, and P. Chambon.** 1988. The N-terminal DNA-binding "zinc finger" of the oestrogen and glucocorticoid receptors determines target gene specificity. EMBO J. **7**:3037–3044.
12. **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.
13. **Hoess, A., A. K. Arthur, G. Wanner, and E. Fanning.** 1988. Recovery of soluble, biologically active recombinant proteins from total bacterial lysates using ion exchange resin. Bio/Technology **6**:1214–1217.

13a. **Höss, A., I. Moarefi, K.-H. Scheidtmann, L. J. Cisek, J. L. Corden, I. Dornreiter, A. K. Arthur, and E. Fanning.** 1990. Altered phosphorylation pattern of simian virus 40 T antigen expressed in insect cells by using a baculovirus vector. J. Virol. **64**:4799–4807.

14. **Klug, A., and D. Rhodes.** 1987. "Zinc fingers": a novel protein motif for nucleic acid recognition. Trends Biochem. Sci. **12**:464–469.
15. **Lanford, R. E.** 1988. Expression of simian virus 40 T antigen in insect cells using a baculovirus expression vector. Virology **169**:72–81.
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. **McVey, D., L. Bizuela, I. Mohr, D. R. Marshak, Y. Gluzman, and D. Beach.** 1989. Phosphorylation of large tumour virus antigen by cdc2 stimulates SV40 DNA replication. Nature (London) **341**:503–507.
18. **McVey, D., M. Strauss, and Y. Gluzman.** 1989. Properties of the DNA-binding domain of the simian virus 40 large T antigen. Mol. Cell. Biol. **9**:5525–5536.
19. **Mohr, I. J., Y. Gluzman, M. P. Fairman, M. Strauss, D. McVey, B. Stillman, and R. D. 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.
20. **Murphy, C. I., B. Weiner, I. Bikel, H. Piwnica-Worms, M. K. Bradley, and D. M. Livingston.** 1988. Purification and functional properties of simian virus 40 large and small T antigens overproduced in insect cells. J. Virol. **62**:2951–2959.
21. **Ryder, K., S. Silver, A. 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.
22. **Ryder, K., E. Vakalopoulou, R. Mertz, I. Mastrangelo, P. Hough, P. Tegtmeyer, and E. Fanning.** 1985. Seventeen base pairs of region I encode a novel tripartite binding signal for SV40 T antigen. Cell **42**:539–548.
23. **Scheidtmann, K.-H., B. Echle, and G. Walter.** 1982. Simian virus large T antigen is phosphorylated at multiple sites clustered in two separate regions. J. Virol. **44**:116–133.
24. **Scheidtmann, K.-H., M. Harding, 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.
25. **Schiff, L. A., M. L. Nibert, and B. N. Fields.** 1988. Characterization of a zinc blotting technique: evidence that a retroviral gag protein binds zinc. Proc. Natl. Acad. Sci. USA **85**:4195–4199.
26. **Schneider, J., and E. Fanning.** 1988. Mutations in the phosphorylation sites of simian virus 40 (SV40) T antigen alter its origin DNA-binding specificity for sites I and II and affect SV40 DNA replication activity. J. Virol. **62**:1598–1605.
27. **Stillman, B.** 1989. Initiation of eukaryotic DNA replication in vitro. Annu. Rev. Cell Biol. **5**:197–245.
28. **Struhl, K.** 1989. Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. Trends Biochem. Sci. **14**:137–140.