Specific Protein Binding to the Simian Virus 40 Enhancer In Vitro

ALAN G. WILDEMAN,[†] MARTIN ZENKE,[‡] CHRISTIAN SCHATZ, MARGUERITE WINTZERITH, THOMAS GRUNDSTRÖM,[§] HANS MATTHES, KEIKICHI TAKAHASHI,^{||} AND PIERRE CHAMRON*

Laboratoire de Génétique Moléculaire des Eucaryotes du Centre National de la Recherche Scientifique, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'Institut National de la Santé et de la Recherche Médicale, Institut de Chimie Biologique, Faculté de Médecine, 67085 Strasbourg-Cédex, France

Received 13 December 1985/Accepted 12 March 1986

HeLa cell nuclear extracts and wild-type or mutated simian virus 40 enhancer DNA were used in DNase I footprinting experiments to study the interaction of putative *trans*-acting factors with the multiple enhancer motifs. We show that these nuclear extracts contain proteins that bind to these motifs. Because point mutations which are detrimental to the activity of a particular enhancer motif in vivo specifically prevent protection of that motif against DNase I digestion in vivo, we suggest that the bound proteins correspond to *trans*-acting factors involved in enhancement of transcription. Using mutants in which the two domains A and B of the simian virus 40 enhancer are either separated by insertion of DNA fragments or inverted with respect to their natural orientation, we also demonstrate that the *trans*-acting factors bind independently to the two domains.

Results of in vitro (12, 13, 16, 19) and in vivo (10, 14) studies have indicated that specific *trans*-acting factors are involved in enhancer function. We have recently demonstrated that the simian virus 40 (SV40) enhancer is composed of several sequence motifs, the association of which is indispensable for the generation of enhancer activity (20). If the *trans*-acting factors were exerting their function by binding to these enhancer motifs, we would expect that this binding would be specifically prevented by mutations which are known to be detrimental to enhancer function in vivo. We performed DNase I footprinting experiments on enhancer DNA by using a HeLa cell nuclear extract and show here that proteins are bound to the various enhancer motifs. Moreover, in agreement with the above expectation, we demonstrate that point mutations which impair the function of a particular enhancer motif in vivo (20) specifically prevent protein binding to that motif in vitro. These results strongly support the conclusion that the SV40 enhancer consists of multiple elements, which bind specific *trans*-acting factors, and that it is the integrated functioning of all of these elements which is responsible for enhancer activity.

(A preliminary report of these results was presented at the meeting on "Eukaryotic Transcription: the Role of Cis- and Trans-Acting Elements in Initiation" [*Current Communications in Molecular Biology*, p. 19–26. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.].)

MATERIALS AND METHODS

DNase I footprinting. All of the recombinants used in this study have been described previously (20). Nuclear extracts of HeLa cells were prepared as described previously (6, 19). Unless otherwise specified, the labeled DNA templates were

obtained as follows. The late-coding and early-coding strands (see Fig. 1 and 3) were labeled by first digesting the SV40 recombinants with *Eco*RI or *Hind*III, respectively, and then phosphorylating these sites with polynucleotide kinase and [γ -³²P]ATP (specific activity, 5,000 Ci/mmol). After subsequent digestion with *Hind*III or *Eco*RI, respectively, the *Hind*III-*Eco*RI fragments were purified on 6% polyacrylamide gels and electroeluted. Unless otherwise specified, footprinting reactions were carried out by preincubating for 15 min on ice the amount of extract indicated in the legends to the figures with 25 ng of pBR322 (linearized at the *Bam*HI site) in a 10- μ l volume containing 10% glycerol, 50 mM KCl, 3 mM MgCl₂, 4 mM spermidine trihydrochloride, and 0.5 mM dithiothreitol. After preincubation, 2 μ l of labeled template DNA (approximately 5 fmol, 3×10^3 to 10×10^3 cpm per reaction) was added, and the samples were gently vortexed and then incubated at 20°C. After 10 min, 2 μ l of a DNase I solution (25 to 75 μ g/ml, with the appropriate amount of DNase I determined for each nuclear extract; Cooper Biomedical, Inc., West Chester, Pa.), freshly diluted in 20% glycerol–1 mM MgCl₂–20 mM KCl–1 mM dithiothreitol from a 1-mg/ml stock solution, was added, and the digestion was allowed to proceed for 90 s at 20°C. Reactions were terminated by sodium dodecyl sulfate (SDS) phenol-CHCl₃ extraction, and prior to ethanol precipitation ammonium acetate was added to a concentration of 1.0 M. The samples were analyzed on an 8% salt gradient polyacrylamide gel as described previously (3).

RESULTS

DNase I footprint over the enhancer in nuclear extracts. The SV40 promoter region was 5' end labeled with ³²P on the early- (E) or late- (L) coding strand at the *Hind*III or *Eco*RI sites of the wild-type recombinant pA0, which contains only one 72-base-pair (bp) sequence (see Fig. 1 and 3; reference 20). When the DNA was incubated with increasing amounts of nuclear extract (1 to 6 μ l) and subsequently treated with DNase I, the pattern of digestion characteristic of naked DNA was very specifically modified (compare lane 0 with lanes 1 to 6, Fig. 1). One of the most striking modifications was the appearance of a DNase I hypersensitive site at position 221 on the late strand (L5, Fig. 1A). The E-coding strand exhibited a similar, albeit less striking, hypersensitive

* Corresponding author.

[†] Present address: Department of Molecular Biology and Genetics, University of Guelph, Guelph, Ontario, Canada N1G 2W1.

[‡] Present address: European Molecular Biology Laboratory, D-6900 Heidelberg, Federal Republic of Germany.

[§] Present address: Unit for Applied Cell and Molecular Biology, University of Umeå, S-901 87 Umeå, Sweden.

^{||} Present address: Medical Institute for Bioregulation, Kyushu University, Maedashi 3-1-1 Higashi-ku, Fukuoka 812, Japan.

FIG. 1. DNase I footprinting over the SV40 promoter region in nuclear extracts. The *HindIII-EcoRI* promoter template was derived from the wild-type recombinant pA0 which contained only one 72-bp sequence (20). The results obtained with the L- and E-coding strands are shown in panels A and B, respectively. DNase I footprinting was carried out as described in the text, with the amount (in microliters) of HeLa cell nuclear extract indicated above each panel. A + G sequence ladders of the labeled template were run in parallel. SV40 nucleotide numbers follow the BBB system (18). The positions (in parentheses) of the enhancer sequence, the 21-bp repeat region with its six GC-rich motifs (I to VI), the TATA box, the early-start sites (EES [17]), and some key restriction enzyme sites are indicated. To the left of each autoradiogram is indicated the regions that have increased (●) or decreased (○) sensitivity to DNase I after incubation in the nuclear extract.

ity at almost the analogous position (nucleotide 223, E3, Fig. 1B). Neighboring nucleotides (position 219 on the L-coding strand and position 225 on the E-coding strand) also exhibited increased sensitivity to DNase I. Additional sites with sensitivity to DNase I increased by incubation with the nuclear extract were located at L3 (position 240) and L7 (position 183) on the L-coding strand and E0 (position 284) and E6 (position 180) on the E-coding strand (a weaker site may be present between E4 and E5 at position 193). On the other hand, several enhancer sequences became protected by incubation with the nuclear extract: L1 (positions 258 to 277), L2 (positions 244 to 250), L4 (positions 226 to 236), and L6 (positions 190 to 218, which may be divided into two subregions: positions 190 to 197 and 199 to 218, according to the kinetics of protection) on the L-coding strand; and E1 (positions 258 to 274), E2 (positions 237 to 248), E4 (positions 199 to 221), and E5 (positions 185 to 191) on the E-coding strand. The amount of extract required to protect regions L4 and E2 was higher than that necessary to protect the other regions of the enhancer and to generate the hypersensitive sites.

The footprint over the 21-bp repeat region (21-bp repeat and 22-bp sequence) was identical to that recently described by using nuclear extracts, in that both the E- and L-coding strands were protected (E7 and L8) and DNase I-hypersensitive sites were observed on the E-coding strand (E8) at the border of the 21-bp repeat and AT-rich regions (position 35) and on the L-coding strand within GC motifs I and IV at positions 44 and 76 (2). As previously noted (2) protection of

GC motif I was less pronounced at low concentrations of extracts than that of the other GC motifs (Fig. 1B).

Proteins are bound to the sequence motifs which are essential for enhancer activity in vivo. The ability of the SV40 enhancer to stimulate early transcription in vitro is impaired by the presence of deletions known to lower its in vivo activity (19). The nucleotide sequences responsible for the enhancer activity in vivo now have been precisely identified by generating a systematic series of clustered point mutations (see Fig. 3; reference 20). The SV40 enhancer which spans approximately 100 bp in its form which contains only one 72-bp sequence appears to be composed of two broad domains, A and B, each of which exhibits little enhancer activity by itself. These domains contain specific sequence motifs which are responsible for their function: the SphI motifs I and II and the P motif for domain A, and the two GT motifs I and II for domain B (see Fig. 3 and reference 20). An enhancer sequence which is simultaneously mutated in positions 202 to 204, 245 to 249, and 268 to 272 is inactive in vivo (mutant pA94 [20]). Similarly, this mutated enhancer sequence is neither able to stimulate SV40 early transcription in vitro nor to compete for the *trans*-acting factor(s) required for enhancer activity (data not shown [19]). We performed competition experiments (data not shown) to test whether the present pattern of DNase I footprinting over the enhancer could be due to the binding of the specific *trans*-acting factor(s) identified in our earlier studies (13, 19). The addition of a truncated pA0 wild-type unlabeled competitor fragment extending from *KpnI* to *BamHI* (see Fig. 1 and 3) to

FIG. 2. DNase I footprinting over SV40 promoter regions containing point-mutated enhancers. Footprinting reactions were performed as described in the text and in the legend to Fig. 1 with 6 μ l of nuclear extract per reaction. The first lane in each panel is the pA0 template, and subsequent lanes are the mutated pA templates, as indicated at the top of each lane (see Fig. 3; for the nucleotide changes present in these mutants, see reference 20). pA12c contains the mutation in the same position as pA12 except that the sequence was changed to ACC (20). The position of each mutation is shown to the right of each panel as a small solid circle, and the location of the regions of DNase I protection and hypersensitivity are indicated as described in the legend to Fig. 1. In control experiments all mutant DNAs were digested with DNase I in the absence of nuclear extract to check that their pattern of digestion was similar to that seen for pA0 (see Fig. 1) (data not shown; data are available on request). Only mutant pA24 showed an additional cut site at the position of the mutation on the L-coding strand of its naked DNA, which resulted in a band which was also present in the presence of extract (see lane pA24 in panel A). The early strand of the pA26 template was overdigested in this particular experiment, but the deprotection in E4 caused by this mutation was identical in other experiments in which the extent of DNase I digestion was lower.

the footprinting reaction decreased the extent of DNase I protection or hypersensitivity over the E- and L-coding strands of the enhancer. By contrast, the addition of the mutated competitor fragment purified from pA94 had no significant effect on the regions of protection or hypersensitivity. On the other hand, neither the protection over regions L8 and E7 nor the hypersensitivity at positions L7, E6, and E8 was lost in the enhancer-specific competition with wild-type competitor fragments. Results of competition and mutation analyses reported elsewhere (2) indicate that they result at least in part from protein binding in the 21-bp repeat region.

These competition experiments support the idea that the DNase I footprint over the enhancer is due to protein(s) which recognize specific sequences. If these proteins were to be involved in the generation of enhancer activity *in vivo*, we would expect that mutations which are known to be detrimental *in vivo* would interfere with the binding of the protein to the enhancer *in vitro* and would specifically alter the DNase I footprinting pattern. In agreement with this expectation, the point mutations which in the pA series resulted in a marked decrease of enhancer activity *in vivo* (20) strikingly affect the DNase I footprinting pattern (see Fig. 2 and 3). The mutation in pA4 led to a complete deprotection throughout E1 and L1, a slight deprotection in E2 and L2, and a disappearance of the moderately hypersensitive site at E0. The mutation in pA12c did not affect L1 and E1 protection but caused a deprotection at L2 and E2, and on the L-coding strand it had the striking effect of eliminating the hypersensitive site at L3, decreasing the protection in the distal part of L4, and lowering the hypersensitivity at L5. On the E-coding strand this mutation also resulted in stronger digestion of nucleotides 234 to 236 and 255, which suggests that the E2 region may extend from nucleotides 234 to 255.

Mutation pA15 caused a strong increase in hypersensitivity at L3 and also a decrease in hypersensitivity at L5; it deprotected one residue in L4 and E2. Mutations in either pA22 or pA26 resulted in general deprotection of E4 and L6 up to position 198, but it is noteworthy that they had no effect on protection of the other regions. However, the hypersensitivity at L5 was strongly diminished by the mutation in pA22, whereas it was moderately decreased in pA26. Although pA32 has very little effect on enhancer activity *in vivo*, there was a deprotection of the residues at that position on both strands, identified on the E-coding strand as region E5 and on the L-coding strand as part of region L6.

In marked contrast, the point mutations pA8, pA20, and pA24, which have little or no effect *in vivo* (Fig. 3) (20) resulted only in marginal alteration of the pattern of protection and hypersensitivity on the E- and L-coding strands. pA20, which is located at the position of hypersensitive sites L5 and E3, caused only a slight decrease in the level of hypersensitivity at E3. pA8 had no detectable effect on the overall DNase I pattern, whereas pA24 caused some deprotection of L6 and E4, but much less efficiently than did the neighboring mutations pA22 and pA26.

All mutated enhancer templates were digested with DNase I in the absence of extract to verify that the naked DNA was digested in the same manner as pA0 (data not shown; data are available on request); only pA24 naked DNA showed a noticeable cut site on the L-coding strand at the position of the mutation, which accounts for the band present in this mutant at that site (Fig. 2). It is important that none of the enhancer mutations affected the protection observed over the 21-bp repeat region (L8 and E7) nor the hypersensitive sites L7 and E6 which were generated by the 21-bp repeat region (see above).

From the results presented above, we conclude that point

FIG. 3. Comparison of the location of regions of DNase I protection and hypersensitivity observed in vitro with the results of the in vivo activity of enhancer point mutants. The E- (upper) and L- (lower) coding strands of the SV40 72-bp and 5'-flanking sequence of pA0 are shown (together with some relevant restriction sites) at the top of the figure (the wild-type sequence at positions 101 to 103 was mutated to generate the *Bam*HI site in pA0, as described previously [20]). The regions of DNase I protection identified in Fig. 1, confirmed by competition experiments (data not shown) and point mutation analysis (Fig. 2), are shown by a solid line. The broken lines flanking region E2 indicate those residues which, based on competition footprinting studies (data not shown), are not clearly protected but based on point mutation footprinting (Fig. 2) appear to be included in this region. Nucleotides that have sensitivity to DNase I that is increased strongly (large arrows) and moderately (small arrows) are indicated. The locations of the SV40 GT motifs I and II (solid lines, I and II), the Sph motifs I and II (double lines, I and II), and the P motif (single line, P) are indicated (see text [20]). Below the enhancer sequence is shown the profile of the in vivo effect of point mutations on transcriptional efficiency of the enhancer (expressed relative to pA0 [20]), and below the profile are shown the nucleotide sequences that are present in each mutated template of the pA series (pA1 to pA34). The location of *Xba*I, *Sal*I, and *Xho*I restriction sites introduced by site-directed mutagenesis (mutants pA211, pA213, and pA204 [20]) and used for the experiments shown in Fig. 4 and 5 are indicated at the bottom of the figure.

mutations which diminish the activity of the enhancer in vivo interfere with the interaction with enhancer-specific protein(s) in vitro. The results of several footprinting, competition, and point mutation experiments are summarized in Fig. 3. The striking observation is that the regions of protection lie directly over those sequences in which point mutations have the largest effect on transcription in vivo, and the hypersensitive sites occur at residues which tolerate the presence of mutations. Furthermore, the SV40 enhancer appears to contain several domains which can bind proteins independently of one another. Nucleotides between approximately positions 180 and 220 (regions L6, E4, and E5) and between positions 225 and 278 (regions L1, L2, L4, E1, and E2) appear to define at least two enhancer domains. No point mutations in one affect binding in the other, and they are separated by a major DNase I cut site (L5 and E3), which becomes more apparent as enhancer-specific factor(s) bind (Fig. 1). Within each of these two domains there is also evidence of independent binding sites (such as the region defined by L1 and E1, the mutation of which [pA4] had little effect on L2 and E2; see Fig. 2 and 3). The following analyses were carried out to determine more precisely how independent the various binding domains are.

Are protein bindings to the protected domains independent? Although it is clear that point mutations present in any one of the regions of protection do not cause appreciable deprotection in the other regions, it is possible that the effects of such

mutations on protein binding are too subtle to reveal cooperativity between the various domains. We therefore examined the effect on DNase I footprinting of inserting nucleotides between the regions of protection, which might be expected to more seriously disrupt the interaction of one region with another. The insertion mutants were constructed by using restriction sites (*Xho*I, *Sal*I) introduced into the enhancer by site-directed mutagenesis (Fig. 3) (20). In Fig. 4 is shown the footprinting analyses of the L-coding strands. pSV β , a control recombinant which differs from pA0 only by the absence of the *Bam*HI site at the 72- to 21-bp repeat region border (20) exhibited the same footprinting pattern as pA0 (compare with Fig. 1A), a result which is in agreement with the observation that its enhancer activity is the wild type in vivo (20). The insertion of four nucleotides at the *Sal*I site present in mutant pA213 (Fig. 3) had no effect on the protection observed over L4 and L6 (mutant pA214; Fig. 4). There was, however, a large decrease in the hypersensitivity at L3 and a selective deprotection of the very proximal portion of L1 (extending from the *Sal*I site to nucleotide 263). This part of L1 was also deprotected by the point mutation in pA4 (Fig. 2). In contrast, the four nucleotides inserted between regions L4 and L5 at the *Xho*I site of mutant pA204 (mutant pA205; Fig. 3) (20) had no effect on the regions of protection or hypersensitivity anywhere in the enhancer, although the hypersensitivity at L5 was extended toward the inserted nucleotides. When a larger insertion (18

FIG. 4. DNase I footprinting of enhancer sequences containing insertions. The pSV β template is identical to that of pA0, except that it possesses the SV40 wild-type sequence at the location of the *Bam*HI site of pA0 (see Fig. 3). pA214, containing a 4-bp insertion at the *Sal*I site of pA213, and pA205 and pA207, containing a 4- and an 18-bp insertion, respectively, at the *Xho*I site of pA204, have been described (20) (see Fig. 3). The positions of the restriction sites *Sal*I and *Xho*I are shown to the left of the wild-type pSV β template (arrows), and to the left of each of the other three templates is shown the position and size of the insertion. Footprinting reactions were performed as described in the legend to Fig. 1 with 6 μ l of nuclear extract, but only results for the L-coding strand are shown here. The regions of DNase I protection and hypersensitivity identified in Fig. 1 are shown to the right of each template. A+G sequence ladders of the labeled templates were run in parallel (left lanes). nt, Nucleotide.

nucleotides) was present at the *Xho*I site (pA207) (20), the hypersensitivity at L3 was unchanged, but most of the inserted nucleotides were cut by DNase I. It is striking that no other change occurred in the footprint protection of any region of the enhancer, except for a slight decrease in hypersensitivity at L3. In vivo, the four-nucleotide insertion at the *Sal*I site reduced the efficiency of the enhancer to 23% that of the wild type, while both the 4- and 18-nucleotide insertions at the *Xho*I site reduced it to approximately 50% (20).

Because none of the insertions described above affected the protection observed in region L6 and because point mutations outside of this region have no effect on L6 and E4 protection, we conclude that there exists an independent binding domain A, which is identified by the in vitro footprinting experiments and the in vivo transcription data (domain A, Fig. 3). Although the identification of distinct elements within the large domain B located upstream of the major hypersensitive sites L5 and E3 is more difficult, our results suggest that there may be two interacting subdomains (B1 and B2) in this part of the SV40 enhancer. One element (B1) may lie in the sequences upstream of nucleotide 260 and encompasses much of the protected regions L1 and E1. The point mutant pA4 prevented binding over this region, whereas the four-nucleotide insertion at the *Sal*I site did not. The insertion (mutant pA214, Fig. 4) caused deprotection only up to nucleotide 263, leaving the remainder of L1 well

protected, yet at the same time resulted in a loss of hypersensitivity at L3, exactly as did the point mutations in pA12c (Fig. 2). In contrast, the deprotective effect of the point mutant pA12c extended primarily throughout the segment located between nucleotides 225 and 255 (Fig. 2; see above), suggesting that this region may constitute a distinct element (B2) within domain B and supporting the notion that most of the sequences in L1 and E1 constitute separate elements. That subdomains B1 and B2 nevertheless may not be fully independent from one another is indicated by the effect of mutation pA4, which caused a slight deprotection in regions L2 and E2.

The possible existence of at least two proteins which bind relatively independently of one another to the SV40 enhancer suggests that it should be possible to invert domains A and B without losing the protein interactions that are responsible for the DNase I footprint. Two restriction sites *Xho*I (pA204) and *Xba*I (pA211) were constructed in pA0 to generate such inversions (20). The creation of the *Xho*I site had no effect on the footprint (compare Fig. 1A and pA204 in Fig. 5), whereas the *Xba*I site (pA211) caused some deprotection at the position of the mutation but had no effect on the regions of protection over the enhancer. Inversion of the enhancer element from *Bam*HI to *Kpn*I (pA301, Fig. 5E) resulted in the same DNase I footprint pattern as that found for the E-coding strand shown in Fig. 1B. The competition with the wild-type enhancer *Kpn*I-*Bam*HI fragment (+pA0)

FIG. 5. DNase I competition footprinting over enhancer regions containing inverted domains. pA211 and pA204 are identical to pA0, except for the presence of the *XbaI* or *XhoI* sites, respectively (20) (see Fig. 3). The remaining three templates had the position of the inverted sequence (*BamHI-XhoI* in pA302, *XhoI-XbaI* in pA303, and *BamHI-KpnI* in pA301) indicated to the left of each panel, with the large arrow showing the orientation of the enhancer sequences with respect to the SV40 early promoter. Footprinting reactions were carried out with labeled *HindIII-EcoRI* fragments, as described in the text and the legend to Fig. 1, with 6 μ l of nuclear extract; but the preincubation on ice was carried out in the presence of 50 ng of linearized pBR322 and either 0 (C) or 50 ng of electrophoretically purified unlabeled *KpnI* to *BamHI* fragments (see Fig. 1 for position of these restriction sites) prepared either from wild-type pA0 (+pA0) or from the pA94 mutant (+pA94), the enhancer of which was transcriptionally inactive because of the clusters of point mutations (see text and reference 20). Only the results for the L-coding strand are shown here. A sample digested without the addition of nuclear (N) extract is included in each case. The regions of DNase I protection and hypersensitivity corresponding to those shown in Fig. 1 are indicated to the right of each panel. Note that whenever a sequence is inverted these regions are those seen on the E-coding strand in Fig. 1. A+G sequence ladders of the labeled templates were run in parallel as indicated.

confirmed that the enhancer-specific factor(s) was recognizing the same sequences regardless of the enhancer orientation in the SV40 promoter region. In both orientations identical deprotection was brought about by wild-type competitor pA0, with region E5 resisting deprotection, whereas the mutated competitor pA94 (see above and legend to Fig. 5) did not lift the footprint (Fig. 5; data not shown). In Fig. 5C and D it is shown that if either domain A (*BamHI* to *XhoI*) or domain B (*XhoI* to *XbaI*) was inverted, both would continue to be protected in a manner identical to that when they were in their natural orientations. The inversions of domain A (pA302) revealed a protection identical to that seen over E4 and E5 in Fig. 1; and the competition with wild-type pA0 resulted in clear deprotection of E4, whereas E5 was not significantly affected. When domain B was inverted (pA303), the protected regions E1 and E2 were specifically competed for by wild-type competitor pA0 but not by mutated competitor pA94. In both pA302 and pA303 constructions, inversion of each domain did not alter the DNase I pattern on the other domains.

These results support the notion that specific protein(s) recognize each of the two large enhancer domains and that binding occurs regardless of their relative orientation.

DISCUSSION

We have shown previously that a specific *trans*-acting factor(s) is involved in stimulation of transcription by the SV40 enhancer in nuclear (19) or whole-cell (13) extracts.

Using a DNase I footprinting assay, we have demonstrated here that specific enhancer sequences are protected from DNase I digestion, when incubated with a nuclear extract (Fig. 1 to 3). Because these protected sequences correspond to the enhancer motifs which were identified *in vivo* as being important for enhancer activity (Fig. 3) (20), we suggest that the proteins which footprint on these sequences are involved in the generation of the enhancer effect both *in vivo* and *in vitro*. It is particularly striking that mutations which are detrimental *in vivo* to the function of a particular enhancer motif lift the *in vitro* footprint only on this particular motif and that sites of DNase I hypersensitivity are located at the boundaries between the motifs, at positions at which mutations have little effect *in vivo*. These sites are presumably not covered by proteins and perhaps have an altered DNA structure (5, 7).

The results of this study support the notion that the enhancer consists of two independent domains A and B, as defined previously (20), because none of the mutations located in domain A affect the footprint on domain B and vice versa. Our data also support our previous proposition (20) that domain B may be composed of two almost independent subdomains. Regions E1 and L1 correspond to the footprint of the protein bound to the upstream GT motif II, whereas the protein bound to the downstream GT motif I footprints on E2 and L2. Because mutations pA12c and pA15 affect both the hypersensitive site L5 and the footprint on L4, it is not clear whether this region corresponds to the

binding site of an additional protein (which may be bound to the motif 5'-TCCCCAG-3' [20]) or represents part of the footprint of the protein bound to the GT motif I. E4 and most of the L6 footprints were positioned on the two Sph motifs. The fact that both pA22 and pA26 mutations lift the footprint in a very similar fashion suggests that there is either a single protein bound to the two motifs or a highly cooperative binding of two proteins. Part of the L6 and E5 regions may correspond to the footprint of a protein on the P motif (Fig. 3) (20). It is worth noting that although mutation pA32 has no apparent phenotype *in vivo* (Fig. 3), it also results in a deprotection of part of this motif between positions 184 and 190 (Fig. 2A and B). This region is also the only one in which the footprint cannot easily compete with the *KpnI-BamHI* fragment of pA0 (Fig. 5; data not shown), which may mean that the corresponding protein is present in large excess. Thus, it appears that there is a minimum of four protein molecules bound to the enhancer: two in domain A and two in domain B. It will be interesting to investigate whether any of these proteins correspond to the Z-DNA binding protein(s) which recently has been isolated from an SV40 minichromosome late in infection (1).

The footprints obtained with the spacing and inversion mutants (Fig. 4 and 5) demonstrate beyond any doubt that domains A and B independently bind proteins *in vitro*. This independence is consistent with the ability of enhancer domains A and B to generate enhancer activity *in vivo*, irrespective of their orientation and, to some extent, their spacing (20). Additional support for this independence of protein binding to domains A and B stems from a footprinting study of a chimeric SV40-polyomavirus enhancer, in which domain A of SV40 is replaced by domain A of polyomavirus (mutant pA411 [20]). In this recombinant DNase I footprints were observed on both SV40 domain B and polyomavirus domain A (data not shown).

The finding that each of the transcriptionally important motifs of the SV40 enhancer binds a protein raises a number of interesting questions. Are these proteins different from one another? In view of the complete lack of sequence homology between the GT and the Sph motifs, it seems probable that they are recognized by different proteins. However, because one can generate enhancer activity not only by associating domains A and B but also, to some extent, by duplication of either domain A or domain B (20), the question arises as to whether enhancer activity can be generated in alternative ways from proteins that possess different functions or whether proteins with similar functions can bind to different sequences. Another interesting question concerns the cell-type specificity of enhancers sharing homologous motifs. For instance, the heavy-chain immunoglobulin enhancer has several enhancer motifs in common with the SV40 enhancer (20, and references therein) and can compete for its effect *in vitro* (13). Nevertheless, it appears to be much more active in lymphoid B cells than in other cells (see references 10 and 15 and references therein), whereas the activity of the SV40 enhancer is ubiquitous. Does this mean that there are families of factors that recognize sequences sharing the same motif, that neighboring sequences are in some way important for the function of a given motif, or both, or that other control mechanisms are operating to maintain cell-type specificity?

It has been reported previously (see references 8, 9, and 11 and references therein) that late in infection a subpopulation of SV40 minichromosomes exhibits a nucleosomal gap over the enhancer region. Moreover, such a gap can be generated *in vivo* over enhancer sequences that are separated from the

rest of the SV40 early promoter and origin regions (9). The results of this study raise the possibility that it is the binding of specific protein factors to the enhancer motifs which prevents the formation of a nucleosome. However, it should be kept in mind that the nucleosomal gap has been observed under conditions in which the SV40 minichromosomes replicate, whereas this study was performed *in vitro* with nuclear extracts of uninfected cells. Studies in which the genomic footprinting technique of Church and Gilbert (4) is used are required to determine whether a protein footprint pattern similar to that found in the present *in vitro* study is also generated over the enhancer region under these *in vivo* replicative conditions.

In summary, we have shown here that the various motifs of the SV40 enhancer do indeed bind specific protein factors to create a large nucleoprotein complex which is presumably involved in the generation of enhancer activity. Results of this study pave the way to the challenging project of purifying these proteins and producing them in abundance to elucidate how this nucleoprotein complex enhances initiation of transcription.

ACKNOWLEDGMENTS

We thank I. Davidson and M. Vigneron for useful discussions and a critical reading of the manuscript, our colleagues for generous gifts of recombinants, and the culture team for growing the cells. We are most grateful to C. Werlé and B. Boulay for the illustrations and to the secretarial staff for patience in typing the manuscript.

This work was supported by grant ATP 6984 from the Centre National de la Recherche Scientifique (CNRS), grant PRC 124026 from the Institut National de la Santé et de la Recherche Médicale (INSERM), a grant from the Fondation pour la Recherche Médicale, grant 82V1283 from the Ministère de la Recherche et de la Technologie, and a grant from the Association pour le Développement de la Recherche sur le Cancer. K. Takahashi, A. Wildeman, T. Grundström, and M. Zenke were recipients of fellowships from INSERM, JSPS, and CNRS; the Natural Sciences and Engineering Research Council of Canada; INSERM, MFR; and the CNRS, Deutsche Forschungsgemeinschaft, respectively.

LITERATURE CITED

1. Azorin, F., and A. Rich. 1985. Isolation of Z-DNA binding proteins from SV40 minichromosomes: evidence for binding to the viral control region. *Cell* 41:365-374.
2. Barrera-Saldana, H., K. Takahashi, M. Vigneron, A. Wildeman, I. Davidson, and P. Chambon. 1985. All six GC-motifs of the SV40 early upstream element contribute to promoter activity *in vivo* and *in vitro*. *EMBO J.* 4:3839-3849.
3. Biggin, M. B., T. J. Gibson, and G. F. Hong. 1983. Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* 80:3963-3965.
4. Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81:1991-1995.
5. Costlow, N. A., J. A. Simon, and J. T. Lis. 1985. A hypersensitive site in hsp70 chromatin requires adjacent, not internal DNA sequence. *Nature (London)* 313:147-149.
6. Dignam, J. D., R. M. Lebowitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 14:1475-1489.
7. Fox, K. R., and M. J. Waring. 1984. DNA structural variations produced by actinomycin and distamycin as revealed by DNAase I footprinting. *Nucleic Acids Res.* 12:9271-9285.
8. Jakobovitz, E. B., S. Bratosin, and Y. Aloni. 1980. A nucleosome-free region in SV40 minichromosomes. *Nature (London)* 285:263-265.
9. Jongstra, J., T. L. Reudelhuber, P. Oudet, C. Benoist, C. B. Chae, J. M. Jeltsch, D. Mathis, and P. Chambon. 1984. Induction of altered chromatin structures by the SV40 enhancer and

- promoter elements. *Nature (London)* **307**:708–714.
10. Mercola, M., J. Gorman, C. Mirell, and K. Calame. 1985. Immunoglobulin heavy-chain enhancer requires one or more tissue-specific factors. *Cell* **36**:403–411.
 11. Saragosti, S., G. Moyne, and M. Yaniv. 1980. Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. *Cell* **20**:65–73.
 12. Sassone-Corsi, P., J. Dougherty, B. Waslyk, and P. Chambon. 1984. Stimulation of *in vitro* transcription from heterologous promoters by the SV40 enhancer. *Proc. Natl. Acad. Sci. USA* **81**:308–312.
 13. Sassone-Corsi, P., A. Wildeman, and P. Chambon. 1985. A trans-acting factor is responsible for the SV40 enhancer activity *in vitro*. *Nature (London)* **313**:458–463.
 14. Schöler, H. R., and P. Gruss. 1984. Specific interaction between enhancer-containing molecules and cellular components. *Cell* **36**:403–411.
 15. Serfling, E., M. Jasin, and W. Schaffner. 1985. Enhancers and eukaryotic gene transcription. *Trends Genet.* **1**:224–230.
 16. Sergeant, A., D. Bohmann, H. Zentgraf, H. Weiher, and W. Keller. 1984. A transcription enhancer acts *in vitro* over distances of hundreds of base-pairs on both circular and linear templates but not on chromatin-reconstituted DNA. *J. Mol. Biol.* **180**:577–600.
 17. Takahashi, K., M. Vigneron, H. Matthes, A. Wildeman, M. Zenke, and P. Chambon. 1986. Stereospecific alignments are required for initiation from the SV40 early promoter. *Nature (London)* **319**:121–126.
 18. Tooze, J. (ed.). 1982. DNA tumor viruses, 9th ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 19. Wildeman, A. G., P. Sassone-Corsi, T. Grundström, M. Zenke, and P. Chambon. 1984. Stimulation of *in vitro* transcription from the SV40 early promoter by the enhancer involves a specific trans-acting factor. *EMBO J.* **3**:3129–3133.
 20. Zenke, M., T. Grundström, H. Matthes, M. Wintzerith, C. Schatz, A. Wildeman, and P. Chambon. 1986. Multiple sequence motifs are involved in SV40 enhancer function. **5**:387–397.