Chromatin Structure of the Simian Virus 40 Late Promoter: a Deletional Analysis

MICHAEL FRIEZ, RAEJEAN HERMANSEN, AND BARRY MILAVETZ*

Department of Biochemistry and Molecular Biology, University of North Dakota School of Medicine, Grand Forks, North Dakota 58202

Received 16 September 1998/Accepted 30 November 1998

The goal of this study was to determine the minimal sequence within the simian virus 40 (SV40) late promoter region, nucleotides (nt) 255 to 424, capable of phasing nucleosomes as measured by its ability to confer the greatest endonuclease sensitivity on adjacent DNA sequences. To identify the minimal sequence, a deletional analysis of the late region was performed by utilizing a SV40 recombinant reporter system. The reporter system consisted of a series of unique restriction sites introduced into SV40 at nt 2666. The unique restriction sites allowed the insertion of test sequences as well as measurement of conferred endonuclease sensitivity. The results of the deletional analysis demonstrated that constructs capable of conferring the greatest nucleosome sensitivities consistently included nt 255 to 280. The activator protein 4 (AP-4) and GTIIC transcription factor binding sequences lie within this region and were analyzed individually. Their abilities to confer nuclease sensitivity upon the reporter nearly matched that of the entire late domain. These results suggest that transcription factors AP-4 and transcription-enhancing factor which binds the GTIIC sequence, alone or in combination, are able to confer significant levels of nuclease sensitivity and are likely involved in the formation of the SV40 nucleosome-free region.

In chromatin, DNA sequences which are either nucleosome-free or have a disrupted nucleosomal structure are also known as nucleosome hypersensitive sites. These sites are thought to be critical in order for trans-acting factors to have access to their cis-acting DNA sequences (12). Nuclease and electron microscopy studies have demonstrated the presence of a nucleosome-free region (NFR) in a fraction of simian virus 40 (SV40) chromosomes found in lytically infected cells (18, 28–31, 35). The information necessary for directing the formation of the NFR lies within the SV40 promoter region, although the mechanisms leading to its formation and maintenance are not very well understood (11, 16, 17, 19, 39). To investigate this further, we have developed an SV40 recombinant reporter system (14) which has allowed us to identify SV40 sequences capable of conferring increased endonuclease sensitivity in chromatin (Fig. 1). The degree of nuclease sensitivity serves as a measure of DNA accessibility and since the reporter sequence does not change, any effects on its nuclease sensitivity must result from changes in chromatin structure caused by the inserted sequence(s) of interest.

Although the SV40 NFR is in a critical region for replication and transcription, its exact function is not completely understood, but it is likely serving as an “open window” for factors required for viral propagation. There is a direct correlation between SV40 chromosomes competent to initiate transcription and those containing a NFR (37, 38). This suggests that transcription factor accessibility is built into the structure of active chromatin and that a number of transcriptional activators aid in preventing the inhibition of transcription by nucleosomes (5, 40). Analysis of the distribution of nucleosomes on SV40 DNA suggests that the location of nucleosomes is neither random nor unique (2). Several studies have indicated that the strongest nucleosome position includes the major initiation start site for late transcription (3, 25, 41). Another study has demonstrated that the strongest nucleosome location is centered at nucleotide (nt) 384 and also includes the major late transcription start site at nt 325 (27). Efficient late transcription proceeds following SV40 replication, therefore it is assumed that one or more mechanisms exist to offset the transcriptional inhibition exerted by this positioned nucleosome.

In a previous report, we utilized our SV40 “reporter” system to identify sequences within the SV40 early promoter and enhancer domains capable of conferring nuclease hypersensitivity (14). In this report, we have performed a similar deletional analysis in order to identify the sequence(s) within the SV40 late promoter capable of conferring increased endonuclease sensitivity. Using this analysis, we have identified specific SV40 late promoter DNA sequences which appear to phase nucleosomes and may be involved in the generation of the SV40 NFR.

MATERIALS AND METHODS

Cells and infections. BSC-1 cells obtained from the American Type Culture Collection were used for the preparation of SV40 reporter viruses and SV40 chromatin. Cells were maintained at 37°C in 5% CO 2 in Eagle’s minimum essential medium (GIBCO) containing 10% fatal bovine serum (GIBCO) and 100 μg of gentamicin (GIBCO) per ml. Subconfluent monolayers of cells were infected with reporter SV40 virus as previously described (21). Infected cells were maintained at 37°C in Eagle’s minimum essential medium containing 2% fatal bovine serum and 100 μg of gentamicin per ml.

Isolation and purification of SV40 chromatin. SV40 reporter chromosomes were isolated from infected nuclei and purified as described previously (24) with modifications. Generally, a single 75-cm 2 T flask was used for each virus. Infected nuclei were extracted with 0.2 ml of nucleus extraction buffer (10 mM HEPES [pH 7.5], 1 mM EDTA, 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride). SV40 reporter chromosomes were separated from virus and cellular debris by sedimentation on a glycerol step gradient containing 1 ml of 10% glycerol in buffer C (10 mM HEPES [pH 7.5], 5 mM KCl, 1 mM EDTA, 0.2 mM MgCl 2, 0.5 mM dithiothreitol, 0.1 mM phenylmethylsulfonyl fluoride) on a cushion of 0.1 ml of 50% glycerol in buffer C. Step gradients were centrifuged in a TLA 100.3 rotor at 50,000 rpm for 35 min in a Beckman TLA 100 ultracentrifuge. Fractions (0.2 ml) were collected from the top to the bottom. The peak of the SV40 chromatin was found in fraction 4.

Restriction endonuclease analysis of SV40 chromatin. Aliquots (20 μl) from fractions 3 and 4 from glycerol step gradients were digested at 37°C for 30 min
with saturating amounts of the appropriate restriction endonuclease following adjustment to reaction conditions with 1/10 volume of buffer A (100 mM Tris-HCl [pH 7.5], 100 mM MgCl₂, 10 mM dithiothreitol, 10 μg of bovine serum albumin per ml). Enzymes used included ApaLI (10,000 U/ml), BglII (8,000 U/ml), MluI (10,000 U/ml), NheI (5,000 U/ml), SalI (20,000 U/ml), and XhoI (20,000 U/ml). All restriction endonucleases were obtained from New England Biolabs.

Agarose gel electrophoresis. Following restriction endonuclease digestions of SV40 chromatin, the products were deproteinized and separated electrophoretically in submerged 1% agarose gels. Gel images were captured by utilizing a UVP GDS8000 Gel Documentation System (Ultra Violet Products). The extent of conversion of form I and II intact SV40 DNA to form III linear SV40 DNA by each endonuclease was quantitated by utilizing Molecular Analyst software (Bio-Rad).

Preparation of reporter constructs. The parental reporter pBM 129 which consists of SV40 mutant strain in or 1411 (a gift from Thomas Shenk [32]) with a polylinker at nt 2666 containing unique restriction endonuclease sites for MluI, ApaLI, PmlI, NciI, and BglII was prepared as previously described (14). DNA sequences from the SV40 promoter were prepared either as PCR products or complementary oligonucleotides and introduced into either the BglII or MluI site adjacent to the reporter. DNA sequences introduced into the BglII site contained a BglII and XhoI site at one end and a BamHI and SmaI site at the other end. Similarly, DNA sequences introduced into the MluI site contained a BssHII and NheI site at one end and a MluI and SalI site at the other end. PCR amplifications were prepared in a thermal cycler (model 480; Perkin-Elmer).

RESULTS

Deletional analysis of the SV40 late promoter with early region present. A fraction of SV40 chromosomes contain a NFR defined by nucleosomes positioned near the ends of the early and late promoter regions. In order to determine whether specific sequences present in the late promoter region act in concert with the early region of SV40 DNA to generate a NFR, constructs containing the early region at one end of the reporter and deletions within the late region at the other end were analyzed for their chromatin structure. Initially, constructs were prepared in which relatively large portions of the late promoter were present. The size and location of each of the deletions are indicated in Fig. 2A. The results of this initial deletional analysis of the late promoter in the presence of an early region are shown in Fig. 2B.

First, a construct containing the early domain introduced at the BglII site, pBM 131-1, was compared to a control construct containing only the reporter. No significant changes in nuclease sensitivity were observed. The introduction of the late promoter into the MluI site, construct pBM 158-39, has a rather profound effect on restriction endonuclease sensitivity, particularly at the MluI and ApaLI sites. Nuclease sensitivity doubles at both the MluI and ApaLI sites compared to those in pBM 131-1, although changes at the BglII and XhoI sites are not apparent. In order to determine which sequences in the late promoter were responsible for this increase in nuclease sensitivity, a series of constructs which contained different regions of the late promoter were prepared. The highest levels of induced nuclease sensitivity were found in constructs containing the regions from nt 255 to 313 (MF4-131) and nt 372 to 424

FIG. 1. Schematic representation of the parental SV40 reporter construct. Both the SV40 promoter (above) and reporter (below) have been expanded to show their organization and relevant restriction endonuclease sites. The major structural elements of the SV40 promoter have been indicated, along with the region in wild-type SV40 chromosomes which is nucleosome free. The early and late domains correspond to approximate positions of the RNA polymerase binding sites for early and late transcription, respectively. The restriction endonuclease sites which are hypersensitive in the wild-type SV40 chromosomes, BglI, KpnI, and NgoMl, are indicated with asterisks. The positions of the deleted copy of the enhancer and the T-antigen intron have also been indicated. The sizes of the various parts of the parental reporter have not been drawn to scale. ORI, origin.
Construct MF4-131 was able to confer the greatest overall nuclease sensitivity in forward and reverse orientations. The region from nt 313 to 372 (MF2-131) had nuclease sensitivities which were consistently 10 to 20% less than the parental construct pBM 158-39 or the constructs containing the other regions (MF3-131 and MF4-131). Since all the constructs tested conferred some increase in nuclease sensitivity at the MluI site, other constructs containing SV40 DNA were ana-
analyzed to exclude the possibility that the conferred increase in nuclease sensitivity was simply a result of the introduction of DNA at this site. The constructs RH3-4 and RH4-2 which contain sequences from nt 123 to 177 and nt 149 to 177, respectively, from the SV40 enhancer were unable to confer significantly elevated nuclease sensitivities on the MluI site or the other adjacent sites in the reporter.

Deletional analysis of MF4-131 in the presence of an early domain. Since the MF4 insert was able to confer the greatest overall nuclease sensitivity and it contains recognition sequences for several DNA binding proteins including AP-1, AP-5, AP-4, transcription-enhancing factor (TEF), and late stimulating factor, it was subjected to a deletional analysis. The size and location of each of the deletions are indicated in Fig.

FIG. 3. Deletional analysis of the MF4 insert in the presence of an early domain. (A) Schematic representations of the reporter and the constructs used in this study. The relative orientation of each domain is indicated. Note that constructs MF13-131 and MF12-131 are in reverse orientation. The black rectangles represent inserted sequences of interest. (B) The percentages of SV40 chromosomes cleaved by digestion with restriction endonuclease as determined by scanning densitometry. The averages are based upon nine or more analyses of SV40 chromosomes from at least three separate infections for each restriction site per construct. Values with an asterisk have been shown to be statistically different ($P < 0.05$) from the value for pBM 131-1 by ANOVA. The data collected for the SalI and NheI sites were compared to initial data from a construct containing an insert with only a SalI and NheI site inserted into the MluI site of pBM 131-1. The initial values for these sites are 33 ± 3 and 32 ± 4 for the SalI and NheI sites, respectively (data not shown).
The results of the deletional analysis of nt 255 to 313 in the presence of an early region are shown in Fig. 3B. The two constructs of this deletional analysis that were able to confer the greatest overall endonuclease sensitivity were MF16-131 and MF22-131. Their abilities to confer nuclease sensitivity were very similar to that of MF4-131. Interestingly, these were the only two constructs which contained nt 255 to 280. Because transcription factor binding sequences GTIIC and AP-4 lie within this region, they were analyzed individually and compared to the other constructs of this deletional analysis. The overall ability of these individual sequences to confer nuclease sensitivity upon the reporter is somewhat less than that of the MF16 and MF22 inserts when they are both present.

During the deletional analysis of the MF4 region, three additional late promoter constructs in the presence of an early domain were also analyzed. The MF5-131, MF6-131, and MF7-131 constructs were further deletions of the initial deletional analysis constructs. The size and location of these sequences are shown in Fig. 4A, and the conferred nuclease sensitivities of these constructs are shown in Fig. 4B and compared to MF4-131-3. The overall nuclease sensitivity conferred by these three constructs is 10 to 25% less than the observed sensitivity in MF4-131-3. The one noticeable exception is construct MF7-131 for which the sensitivity at the MluI and SalI sites is slightly greater than the nuclease sensitivity at the corresponding sites in MF4-131-3.
Analysis of the TFIID transcription factor recognition sequence and the major late transcription start site in the presence of an early domain. The individual recognition sequences for transcription factor TFIID and the major late transcription start sites were analyzed for their abilities to confer endonuclease sensitivity in the presence of an early domain. These sequences were analyzed individually as well as in combination. The location, size, and data collected from these constructs are compared to those of MF4-131-3 and are shown in Fig. 4. Construct MF8-131 contains the TFIID binding site, and MF9-131 contains the major late transcription start site which is found at nt 325. The overall abilities of these sequences to confer nuclease sensitivity do not equal that of MF4-131-3. In fact, nuclease sensitivities are 10 to 20% less at a majority of the reporter restriction endonuclease sites. Construct MF10-131 is a combination of the two individual sites and confers only slightly higher nuclease sensitivity than the individual sequences do.

Analysis of late promoter constructs in the absence of an early domain. SV40 late promoter sequences were also analyzed in the absence of an early domain in order to determine how they function individually. The late promoter sequences which were analyzed are shown schematically in Fig. 5A. The inserts used in this study were the same as those used in the previous studies which included the presence of an early domain. These constructs were compared to parental construct pBM 129-1 which does not contain an insert in the reporter, and the data collected are shown in Fig. 5B. The introduction of the late promoter into the MluI site, construct pBM 165-39, has a rather profound effect on restriction endonuclease sensitivity particularly at the MluI and ApaLI sites. Nuclease sensitivity increases 21 and 26%, respectively for the MluI and the ApaLI sites compared to those in pBM 129-1, although no significant change at the BglII site is apparent. Next, the four inserts of the initial deletional analysis, MF1, MF2, MF3, and MF4 were analyzed in order to determine their ability to confer nuclease sensitivity in the absence of an early domain. Construct MF3-129 was the insert which was able to confer the greatest amount of nuclease sensitivity on the reporter, and the amount of digestion at each site was nearly identical to that of construct pBM 129-1.

Inserts MF16 and MF22 had previously demonstrated the ability to confer nuclease sensitivity in the presence of the early domain. In order to determine their abilities in the absence of the early region, they were individually inserted into the reporter. Both constructs produced results similar to those of parental construct pBM 165-39. In addition to these constructs, transcription factor recognition sequences GTTIC and AP-4 located within MF16 and MF22 were also analyzed. These individual sequences were also able to confer nuclease sensitivities very similar to that of pBM 165-39. The only site which varied by more than 5% in comparison to pBM 165-39 was the Sall site in the GTTIC-129 construct.

Similar levels of conferred nuclease sensitivity were obtained for constructs containing a GTTIC or AP-4 site in the presence or absence of the early region except at the ApaLI site where an increase of approximately 10% was observed when the early sequence was absent (compare Fig. 3B and 5B). This difference in sensitivity at the ApaLI site was also observed when the corresponding parental constructs pBM158-39 to pBM165-39 were compared (compare Fig. 2B and 5B). The basis for this difference is not clear, but we have observed a similar result for the XhoI site when comparing constructs containing an early region phasing sequence in the presence and absence of the whole enhancer (14).

**DISCUSSION**

“Fine tuning” local chromatin structure, particularly in promoter regions of activated genes, is a prerequisite for efficient transcription (33). In this study, our goal was to identify the minimal sequence of the SV40 late promoter capable of conferring the greatest endonuclease sensitivity which would be indicative of local chromatin remodeling. There have not been any previous studies suggesting that transcription factors AP-4 and TEF-1 might be involved in modulating chromatin remodeling events. Most studies that involve AP-4 have addressed only its involvement in transcriptional activation (23). TEF-1 has also been recognized for its involvement in transcriptional activation, particularly in SV40 (8). Interestingly, both AP-4 and TEF-1 participate in transcriptional activation by direct protein interactions; for example, TEF-1 mediates SV40 late promoter transcriptional activation by large T antigen (13). The results from these individual sequences suggest that there may be certain transcription factor binding sites which may be strategically located in regions where chromatin remodeling is required for efficient transcription initiation. This is in agreement with previous studies that have demonstrated that disruption occurring when one factor binds a nucleosome can potenti ate the binding of another factor that would otherwise bind its site poorly (1). Other studies have indicated that chromatin-disrupting sequences are often found with other transcription factor binding sites, which allows the corresponding protein factors to bind more efficiently and with greater affinity (20). Therefore, from the results of this study, it seems likely that the sequence involved in phasing nucleosomes away from the promoter region is positioned at a very strategic location. This location appears to be even more logical considering that the TFIID site is nearly adjacent to this sequence and TFIID cannot bind its target site in a chromatin environment (22).

Whether protein factors are directly responsible for nucleosome phasing remains unclear in this system. If transcription factors are directly involved, then it is certain that accessibility to their target sequences is critical (reference 26 and references therein). It is clear that Pho4 and activated glucocorticoid receptors acting as transcriptional factors must have access to DNA to exert their effects. For example, activated glucocorticoid receptors have the ability to bind nucleosomal DNA (7). Transcription factor Pho4 is different because it requires its initial binding site to be in an 80-bp nucleosome hypersensitive site (34). In SV40, the strong positioning of a nucleosome centered at nt 384 would suggest that the region implicated in conferring nucleosome hypersensitivity usually lies in a nucleosome-sensitive nucleosome linker region (27). This may be crucial for allowing factors the initial access to DNA that ultimately leads to transcriptional activation of the late promoter. Both PHO5 and mouse mammary tumor virus are replication-independent, differing from SV40 late transcription which does require replication. Therefore, this strongly suggests that any model of SV40 late promoter chromatin remodeling needs to include replication as a factor. Furthermore, the possibility of a chromatin remodeling complex, such as SWI-SNF, being involved has not been ruled out.

It has been stated that transcription factors are able to serve multiple functions (10). For example, a transcription factor may be involved in the stabilization of the transcriptional initiation complex and at the same time be responsible for recruiting additional factors to enhance transcription. Transcription factor AP-4 appears to be another example of a protein that is likely to be involved in a variety of different processes that lead to greater levels of transcription. AP-4 can stimulate
in vitro transcription from a nonchromatin template, so it does not seem likely that AP-4 is strictly involved in chromatin remodeling (23). AP-4 and transcription factor AP-1 act in concert to activate SV40 late transcription, and AP-1 has been previously shown to efficiently phase nucleosomes (14). In addition, AP-4 has been shown to contain a number of elaborate dimerization domains, which suggests that it is possible for AP-4 to interact with a wide variety of additional factors (15). AP-4 may in fact be responsible for recruiting a chromatin remodeling complex such as SWI-SNF to the SV40 late promoter. A closer look at AP-4 indicates that there is considerable homology between AP-4 and other transcription factors such as GCN4, Myc, and Max. Recent work suggests that histone acetyltransferase A is targeted to promoter regions by the acidic domain of transcription factors such as GCN4 (6). This suggests the possibility that AP-4 may also have the ability to recruit acetyltransferases to the SV40 late promoter as well as to other promoters with which it is involved. Furthermore, transcription factors Myc and Max are able to bind their target sequences even when they are incorporated into nucleosomes.
11. Suggest that SV40 nt 255 to 280 do in fact confer nuclease-dependent mechanism (4). Initial studies in our laboratory transcription factor PPR1, which functions in a replication-dependent way. Future generations. This model is very similar to the activation has been demonstrated in a number of studies. Furthermore, some, there would be an increase in nuclease sensitivity, which together would lead to inhibition of nucleosome formation over the DNA. AP-4 could then recruit additional factors that successfully compete for DNA access will likely recruit additional factors that result in a more stable transcriptional complex.

When a nucleosome is positioned over the SV40 late promoter, its removal or disruption is required in order for efficient late transcription (2, 27). Our working model proposes that prior to replication, the nucleosome positioned over the late promoter inhibits transcription by preventing transcription factors that are unable to bind nucleosomal DNA access to their binding sites. Viral replication would be expected to temporally create a “more open” DNA template that would allow transcription factors such as AP-4 access to their target sites in the DNA. AP-4 could then recruit additional factors that together would lead to inhibition of nucleosome formation over the promoter region. Due to the lack of a positioned nucleosome, there would be an increase in nuclease sensitivity, which has been demonstrated in a number of studies. Furthermore, the transcription complex could be heritable and passed on to future generations. This model is very similar to the activation of the URA3 gene in Saccharomyces cerevisiae telomerases by transcription factor PPR1, which functions in a replication-dependent mechanism (4). Initial studies in our laboratory suggest that SV40 nt 255 to 280 do in fact confer nucleosome sensitivity only in the presence of replication.

ACKNOWLEDGMENT

This project was supported by NSF EPSCOR grant OSR 9108770 to the State of North Dakota.

REFERENCES


Vol. 73, 1999 CHROMATIN STRUCTURE OF THE SV40 LATE PROMOTER 1997