Sequence and structural requirements for primase initiation in the SV40 origin of replication

---

B.Y.Tseng* and C.E.Prussak

Eukaryotic Regulatory Biology Program, M-103G, School of Medicine, University of California
San Diego, La Jolla, CA 92093, USA

Received November 9, 1988; Revised and Accepted February 6, 1989

ABSTRACT

Primase synthesizes decaribonucleotides for priming of lagging and possibly leading strand synthesis at a replication fork. The sites of initiation by purified mouse primase were shown to be highly specific within the SV40 origin of replication. This study further examines the role of the 27-bp inverted repeat in the origin for initiation. A site is observed on the L-strand template at nucleotide position (np) 22 positioned a similar distance from the 27-bp inverted repeat as sites previously reported on the E-strand. The initiations adjacent to the 27-bp repeat have a higher K_m for rATP than other sites. A deletion within the inverted repeat eliminated initiation at sites proximal to the hairpin on both E and L strands but had no effect at more distant sites. A deletion mutant which left the inverted repeat intact but deleted the initiation sites at np 5210-5220 on the E-strand was not active as a template for proximal sites. These results indicate that primase has two modes of recognition, one that requires the SV40 inverted repeat structure and a specific sequence and another that requires sequence alone. Additional regions of the SV40 genome have also been examined and of approximately 2000 nucleotides of single stranded template examined, only one additional site was observed at np 2412 on the E-strand. This indicates that primase initiations are highly specific for the SV40 origin and their potential functional role is discussed.

INTRODUCTION

The initiation of DNA synthesis at a origin of replication requires multiple steps, one of which is the synthesis of a RNA primer for DNA polymerase (1). For SV40 replication, the RNA primers found on nascent fragments isolated from replicative intermediates have been shown to span the SV40 origin of replication and are decaribonucleotides (2). This size is characteristic of primers found on Okazaki fragments (3-5) and also characteristic of RNA synthesized by eukaryotic primase (6-8). With purified mouse cell primase and single-stranded DNA templates of the SV40 origin, site specific initiation has been demonstrated (9,10). These results suggest that primase, characterized for synthesis of primers of discontinuous DNA synthesis (3-5), may also be required for RNA primer initiation of leading

strand synthesis at an origin of replication. To characterize the in vitro specificity of primase for the SV40 origin of replication we have extended our analysis of SV40 initiation sites by primase.

We had previously observed specific initiation within the SV40 origin of replication at np 5145-5147 and np 5210-5220 on the early (E) strand and in the six G/C box repeats on the late (L) strand as templates (9). The initiations at np 5210-5220 were specifically affected by a deletion, DL1, within the 27-bp inverted repeat that also inactivated origin function in vivo (11) whereas the other sites were not affected. Since the inverted repeat would be expected to function in a symmetrical fashion, we have examined the late side of the origin for additional primase initiation sites. We have further characterized the specificity of the initiation reaction with deletion mutants and have examined other regions of the SV40 genome for potential initiation sites.

METHODS

Replication run-off mapping of initiation sites

Primase (8) and DNA polymerase α (12) were purified from mouse hybridoma cells as described previously. Isolation of single-stranded linear insert DNA by restriction enzyme digestion of the insert from single stranded M13 mp7 DNA has been described (9). Mapping of initiation sites was carried out by replication run-off assay following previous procedures (9). Briefly, primase was incubated with single stranded linear DNA template along with DNA polymerase α , rNTP, and ^{32}P -dNTP. To ensure all initiated products were extended to the end of the template, a second short incubation with *E. coli* DNA polymerase I large fragment was carried out. The initiation sites of the labeled products were determined by electrophoresis along with a sequence ladder of the product strand.

Construction of deletion mutants

Deletion mutants were constructed with the SV40 RII G fragment (blunt end) inserted into the HincII site of M13 mp7 and oriented with the E-strand in the phage DNA. mp7 RIIG-E RF was cut at the single HindIII site (np 5171 in SV40) and deletions made by incubation with Bal31 nuclease (13). Termini were made blunt with S1 nuclease and ligated. Individual phage plaques were screened for the direction and approximate end point of the deletion by annealing 5' or 3' ^{32}P -L strand to individual phage DNA and digesting with S1 nuclease. The phages with the appropriate deletions were picked, sequenced through the deleted region to determine the end points of the deletion, and used for isolating mutant linear single stranded insert DNA.

Figure 1 rATP effects on replication run-off reactions with L-strand template. The L-strand of the SV40 RII G fragment was incubated in a replication run-off reaction as described (9). rATP concentration was varied with the indicated concentrations (0, 0.02, 0.05, 0.1, 0.2, 0.5, and 1 mM). The reaction products were electrophoresed on a 6% polyacrylamide sequencing gel (16) and autoradiographed. The nucleotide positions on the SV40 genome are indicated on the right along with the AT and G/C boxes (only four are shown), located by a chemical degradation sequence ladder (16) of the product strand.

RESULTS

Initiations within the origin of replication

Previously, we found that using the L-strand of the Eco RII G fragment of SV40 which contains the origin of replication (11,14,15) as a single-stranded template, initiation occurred within each of the six G/C repeats of 3'-CCCGCC-5' (9) and were not affected by a 6-bp deletion within the 27-bp inverted repeat (T-antigen binding site II) of the SV40 origin. On the E-strand, initiations at nucleotide positions (np) 5210-5220 were affected by the deletion mutant. If the inverted repeat sequence is involved in primase initiation, it would be expected to have a symmetrical effect. We re-examined initiations on the L-strand template. Initiation at np 22 occurred when rATP concentrations reached 0.5 mM whereas initiations within the G/C boxes occurred at lower concentrations (Fig. 1). This may suggest that in the presence of six repeats of the G/C box, primase is limited for use of the np 22 start or that the high rATP concentration affects a change in site specificity of the enzyme. Increasing primase amounts had little effect on initiation at np 22 with 0.5 mM rATP. Initiations within the 3'-CCCGCC-5' boxes also changed from the 3' side of the G to 5' side, with increasing rATP, although this occurred at lower concentrations. This is probably due to the T residues between the G/C boxes that would re-

Figure 2 rATP effect on E-strand products. The E-strand of SV40 RIIG fragment was incubated in a replication run-off reaction with increasing concentration of rATP. The products were electrophoresed on a 6% polyacrylamide sequencing gel. Only the initiations at np 5210-5220 are shown. No variation of initiation was seen at np 5145-47 with high rATP, similar to what is shown at np 5210 and np 5214.

quire rATP for productive initiation and, as shown below, primers would be too short if rATP is limiting.

To examine if rATP has an effect on the E-strand template, initiations at np 5210-5220 was examined. When primase was incubated with increasing rATP and E-strand template, initiations at np 5218 and np 5220 occurred only at high rATP concentrations whereas initiations at adjacent sites (np 5210, 5214) (Fig. 2) or distal sites (np 5145, 5147) occurred at lower rATP concentration. Previous measurements with primase and simple sequence synthetic templates indicated a K_m for rATP of 150 μ M (8) and this value appears to be consistent with initiation observed at np 5210 and 5214. However, initiations at np 5218 and 5220 require higher concentrations, apparently similar to initiation on the L-strand at np 22 and suggests that these sites may have different initiation requirements.

We had previously found (9) that a 6-bp deletion within the inverted repeat specifically affects the proximal initiation sites at np 5210-5220 but not those at np 5145-5147 on the E-strand (Fig. 3). We used the L-strand as template, at a rATP concentration (1 mM) that sustains primase initiation at np 22, to determine the effect of the DL1 deletion within the inverted repeat. When wild type template is used, starts at np 22 and within the G/C boxes are observed as described above. With the DL1 mutant (16), which has 6-bp deleted between np 5237-5242 in the 27-bp inverted repeat and is functionally inactive as an origin, was used as template only initiations at np 22 are eliminated whereas initiation at the G/C boxes were unaffected (Fig. 3). The consequence of the DL1 deletion is then symmetrical. Both of the regions that are sensitive to the DL1 mutation are approximately the same distance from the 27-bp inverted repeat (Fig. 4A). The 27-bp sequence is a perfect

Figure 3 Effect of the DL1 mutation on primase initiations. Replication run-off assay on the E- or L-strand with wild type (wt) or DL1 mutant template, which has 6-bp deleted at np 5237-5242, were carried out with rATP at 1 mM. The products were run on a 6% polyacrylamide sequencing gel. The early strand reaction is as previously reported (9).

inverted repeat and would be predicted to form a hairpin structure in solution when present as single stranded DNA. These results would suggest that the predicted secondary structure, possibly the sequence of the inverted repeat, may position the enzyme (Fig. 4B).

Structural requirements for initiation

To address whether the inverted repeat would be sufficient as a promoter for primase, we constructed deletion mutants from the HindIII site (np 5171) toward the origin. Several mutants dl 10-11 (np 5170-5182), dl 10-12 (np 5170-5193) and dl 10-13 (np 5170-5201) which deleted sequences downstream of initiation sites at np 5210-5220 had little or no effect on initiations at np 5210-5220 or at np 5145-5147, except to shorten the product length from the np 5210-5220 sites (Fig. 5). With dl 10-41 which deleted np 5161-5229 but left the inverted repeat intact, no initiation was observed from the region immediately adjacent to the inverted repeat (expected positions indicated by arrow) but initiation at the distal start site, np 5145-5147, was not affected

Figure 4 Initiation sites on the SV40 genome.

A. Primase initiation sites in the SV40 origin are indicated by the arrows with the height of the arrow proportional to the frequency of initiation and the direction of the arrow indicating the direction of synthesis. The 65-bp minimal origin of replication (14-16) is indicated as are T-antigen binding sites and the 21 bp repeats containing the six G/C boxes. Not shown are the initiations at np 5145 and 5147. B. Initiation sites are indicated with the expected structure of the 27-bp inverted repeat sequence when the template strand is present as a single stranded template. The 6-bp deleted in the DL1 mutation are indicated by the brackets.

by the dl 10-41 mutation (Fig. 5). This indicates that the inverted repeat or hairpin structure itself is not sufficient for primase initiations but that the nucleotide sequences in np 5210-5220 are also important.

Specificity of initiation

Primase initiation within the 311 bp EcoRII G fragment of the origin of repli-

Figure 5 Effect of deletions downstream of the 27-bp inverted repeat on primase initiations. Unidirectional deletions from the HindIII site (np 5171) toward the origin were constructed as described in Methods. Mutant E-strand templates were incubated in a replication run-off reaction and the products electrophoresed and visualized by autoradiography. The deleted regions were dl 10-11 (np 5170-5182), dl 10-12 (np 5170-5193), dl 10-13 (np 5170-5201), and dl 10-41 (np 5161-5229). The length of the products from initiations at np 5210-5220 are shortened due to the deletions. The expected size of products for dl 10-41 if the 27-bp inverted repeat could function as a promoter is indicated by the arrow. The initiations of np 5145-5147 are indicated.

cation is highly specific. To determine if other regions of SV40 have initiation sites, we have examined Hind III fragments D, E, and F that lie approximately one quarter of the way around the genome from the origin on the early side (fragment D) and on the late side (fragments E and F). Both strands of each fragment were cloned, isolated, and used as templates for primase replication run-off assays. Of the six strands only one, the E-strand of Hind III, showed any initiation sites and it was located at np 2412. Thus, of approximately 2000 nucleotides of SV40 DNA outside the origin region, only one site served as a site of initiation. This indicates the specificity of primase initiations for the SV40 origin.

Minimal primer length

One of the distinct characteristics of eukaryotic primases is their synthesis of a relatively discrete product of 10 ribonucleotides in length (3,4). The size of the products synthesized by primase have also been characterized with the SV40 template as a relatively discrete size of 10 nucleotides (9). To assess what may be the

Figure 6 Minimal primer length required for productive initiation. The E-strand was incubated in an replication run-off reaction with the indicated rNTPs. The reaction products were separated by gel electrophoresis and visualized by autoradiography. Only the reaction products at np 5210-5220 are shown and the position of each start site indicated on the side of the figure.

minimum size of a primer, we used the specific initiations within the SV40 origin to examine which nucleotides are required for productive priming. When combinations of rNTP were added only certain of the starts at np 5210-5220 were utilized (Fig. 6). rCTP is not required and from the nucleotide sequence (Fig. 4) the closest G in the template from any initiation site is 7 nucleotides away. It can be deduced from the sequence what the primer length has to be for successful initiation. The length of the RNA that can be synthesized when only rATP and rGTP are added would be a 5-mer from np 5220, 3-mer from np 5218, 2-mer from np 5214, and 6-mer from np 5210. The intensity of replication run-off at each position indicated that initiation at np 5210 with a primer size of 6 nucleotides is sufficient for full level of synthesis (Fig. 6). With a primer size of 5, at np 5220, there is a reduction to one-third the number of starts, and no initiations were observed from np 5214 and 5218 for primers 2 or 3 nucleotides long. Since the frequency of initiation is reduced when the primer size is 5, the minimum primer length would presumably be 4 or 5 nucleotides.

It has been found that primers synthesized in isolated nuclei misincorporate both deoxynucleotides and ribonucleotides when rNTPs are limiting (17,18). However, with purified primase, the substitutions do not appear to occur with any great frequency as seen by the specificity of the initiation reaction. We have measured the incorporation of deoxynucleotides by primase *in vitro* with synthetic single stranded templates poly (dGdT)_n or poly (dAdC)_n. The incorporation of rCTP versus dCTP and rUTP versus dTTP was made and in both cases, the deoxynucleotide was incorporated at 10% the amount as the ribonucleotides when incubated independently. In addition, deoxynucleotide incorporation was internal

as shown by the production of labeled dinucleotides after alkaline hydrolysis of the products. This level of deoxynucleotide incorporation is apparently not sufficient to allow productive priming from the SV40 sites when the cognate rNTP is omitted.

DISCUSSION

We describe here the highly specific initiation by purified mouse primase for initiation on the SV40 genome. We had previously observed that initiation occurred at specific sites within the 65 bp minimal origin of replication. We examined an additional 2000 nucleotides for initiation sites and have found only one site outside the origin whereas near or within the origin there are essentially 4 regions: on the E-strand at np 5145-5147 and at np 5210-5220 and on the L-strand at np 22 and at each of the six G/C repeats. The SV40 origin is then highly enriched for primase initiations. Other laboratories have carried out analysis of initiation sites with less purified enzymes of primase/DNA polymerase α complex and have found additional starts on the SV40 template (10,19). Our results indicate a high specificity of initiation with purified primase which may suggest that additional protein(s) relax the specificity for recognition sites as might be anticipated for its role in lagging strand priming. The specific sites of initiation are located near or immediately adjacent to the boundaries of the minimal origin. We suggest that their orientation and specificity would indicate that these are sites for leading strand initiation. Additional factors are not essential for the specific priming, but rather, the high specificity observed may be a requirement for leading strand initiation.

The sites we observed with purified primase may also have *in vivo* relevance. In a study of the RNA-DNA junctions of SV40 replicative intermediates, Hay and DePamphilis (2) described a bidirection transition point of leading and lagging strand synthesis on both strands at np 5210. This would be at the transition point between RNA and DNA for initiation events observed on the E-strand from np 5210-5220, although sites on the L-strand at this position are not observed. Also, in studies of fork movement in SV40 replicating intermediates by electron microscopy, Tack and Proctor (20) described three classes of SV40 replicating chromosomes, one class exhibits symmetrical bidirectional fork movement and two others asymmetrical fork movement rates. The symmetrical class constitutes the largest proportion (67%) of the replicating chromosomes (20). The other classes replicate asymmetrically, the distance traveled by the replication forks either biased toward the early side or toward the late side. There is a close correspondence of the origins of these classes of replicating intermediates with primase initiation sites. The symmetrical class origin (np 5203 $\pm$ 27) is at or near np 5210-5220 primase initiation sites which require the inverted repeat sequence. The asymmetric classes also start near

primase initiation sites. The asymmetric early side origin (np 5148±24) is near np 5145-5147. The asymmetric late side class origin (np 29±60) is near the site at np 22. In addition, the strongest of the initiations in the six G/C boxes are immediately adjacent to this site at np 46 and np 57. The correspondence that is observed would suggest that the asymmetric replicating population is initiated on the leading strand by primase.

In addition, studies with single point mutants in the SV40 minimal origin of replication have characterized three domains essential for replication: at np 5211-5220, at T-antigen binding site II, and at the AT sequence (21). Recent studies of T-antigen SV40 origin DNA unwinding activity indicated that AT region mutants affect T-antigen unwinding as do mutants in np 5211-5215 (22). The functional requirement for sequences at np 5216-5220 cannot be accounted for as yet but correspond to the primase initiation sites with a high K_m for rATP at np 5218 and 5220. We suggest that the genetic requirement for these sequences is for primase initiations.

The nucleotide sequences recognized by primase have not yielded any particular consensus, although a characteristic is the high purine content (> 70%) of the primers. However, as we have shown here and from previous studies, the 27-bp inverted repeat within the minimal origin is required for initiations at np 5210-5220 and at np 22, whereas initiation at np 5145-5147 and the six G/C repeats do not depend on the inverted repeat. Also, the inverted repeat structure by itself is not sufficient for initiation since deletion of adjacent sequences are not active as templates. These observations would suggest that there are two modes of template recognition by primase, one requiring sequence alone and the other requiring a 2° structure and sequence. Paradoxically, primase does not exhibit strong specificity when presented with simple single stranded DNA templates. We have found that poly dIT, (a random co-polymer), poly dT, poly (dGT)_n, poly (dAC)_n, and several others tested are very active for oligoRNA synthesis in contrast to natural DNA, such as fd or SV40 single strands (8). The determining factors for primase specificity appear to be more complex than a definable sequence.

ACKNOWLEDGEMENTS

This research was supported by USPHS Grants GM 29091 and T32HL07107.

*To whom correspondence should be addressed

REFERENCES

1. Kornberg, A. (1980). DNA Replication, W.H. Freeman, San Francisco.
2. Hay, R.T. and DePamphilis, M.L. (1982) Cell 28, 767-779.

3. Reichard, P., Eliasson, R. and Soderman, G. (1974) *Proc. Natl. Acad. Sci. USA* 71, 4901-4905.
4. Tseng, B.Y. and Goulian, M. (1977) *Cell* 12, 483-489.
5. Tseng, B.Y., Erickson, J.M. and Goulian, M. (1979) *J. Mol. Biol.* 129, 531-545.
6. Conaway, R.C. and Lehman, I.R. (1982) *Proc. Natl. Acad. Sci. USA* 79, 4585-4588.
7. Tseng, B.Y. and Ahlem, C.N. (1982) *J. Biol. Chem.* 257, 7280-7283.
8. Tseng, B.Y. and Ahlem, C.N. (1983) *J. Biol. Chem.* 258, 9845-9849.
9. Tseng, B.Y. and Ahlem, C.N. (1984) *Proc. Natl. Acad. Sci. USA* 81, 2342-2346.
10. Yamaguchi, M., Hendrickson, E.A. and DePamphilis, M.L. (1985) *Mol. Cell. Biol.* 5, 1170-1183.
11. Myers, R.M. and Tjian, R. (1980) *Proc. Natl. Acad. Sci. USA* 77, 6491-6495.
12. Prussak, C.E. and Tseng, B.Y. (1987) *J. Biol. Chem.* 262, 6018-6022.
13. Legerski, R.J., Hodnett, J.L. and Gray, H.B. (1978) *Nucleic Acids Res.* 5, 1445-1464.
14. DiMaio, D. and Nathans, D. (1980) *J. Mol. Biol.* 140, 129-142.
15. Bergsma, D.J., Olive, D.M., Hartzell, S.W. and Subramanian, K.N. (1982) *Proc. Natl. Acad. Sci. USA* 79, 381-385.
16. Maxam, A. and Gilbert, W. (1980) in *Methods in Enzymology*, vol. 65, Grossman, L. and Moldave, K. Eds. pp. 499-560, Academic Press, New York.
17. Eliasson, R. and Reichard, P. (1978) *Nature*. 272, 184-185.
18. Tseng, B.Y. and Goulian, M. (1980) *J. Biol. Chem.* 255, 2062-2067.
19. Vishwanatha, J.K., Yamaguchi, M., DePamphilis, M.L. and Baril, E.F. (1986) *Nucleic. Acids. Res.* 14, 7305-7323.
20. Tack, L.C. and Proctor, G.N. (1987) *J. Biol. Chem.* 262, 6339-6349.
21. Deb, S., DeLucia, A.L., Baur, C-P., Koff, A. and Tegtmeyer, P. (1986) *Mol. Cell. Biol.* 6, 1663-1670.
22. Dean, F.B., Borowiec, J.A., Ishimi, Y., Deb, S., Tegtmeyer, P. and Hurwitz, J. (1987) *Proc. Natl. Acad. Sci. USA* 84, 8267-8271.