Volume 12 Number 7 1984 Nucleic Acids Research

A detailed analysis of an HSV-1 early promoter: sequences involved in trans-activation by viral immediate-early gene products are not early-gene specific

R.D.Everett

Medical Research Council Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, UK

Received 2 February 1984; Accepted 7 March 1984

ABSTRACT

This report describes a detailed analysis of the functional DNA sequences within the HSV-1 glycoprotein D gene promoter. The transcriptional activity of deletion and insertion promoter mutants was studied after both trans activation, mediated by viral products, and cis activation by a linked SV40 enhancer. Two G-rich areas (upstream of a TATA signal) were identified as important regions of the promoter. These "upstream" signals were active in both orientations. A functional TATA-box region was detected. A second region, not homologous to the concensus TATA sequence, also appeared to have a role in the positioning of the RNA cap-sites, which included both purine and pyrimidine 5'ends. Deletion of the cap-site region resulted in a moderate reduction in transcription. All the promoter elements were important for both cis and trans activated transcription. No sequence specific for viral (trans) regulation was detected, implying that Early promoters are not distinguished by specific sequences. Since HSV-1 and some other animal viruses can activate transcription from unrelated promoters, this process is probably non-specific and applicable to many, particularly extra-chromosomal, genes. The possible mechanisms of this activation are discussed.

INTRODUCTION

The transcriptional programme of Herpes simplex virus type 1 (HSV-1) during lytic infection involves the coordinate control of at least 3 groups of genes, referred to as Immediate-Early (IE or α), Early (E or β) and Late (L or γ) (1,2,3). The IE-genes are transcribed first. Their promoters appear to contain an enhancer element that responds to cellular factors but whose activity is greatly increased by a component of the virus particle (4-8). The functions of the IE-gene products are in most cases unknown, but it is clear that Vmw175, encoded by IE-gene 3, is essential for activation and continuing function of the Early promoters (9,10,11). This paper describes the location of DNA sequences within the glycoprotein D (gD)

Early gene promoter which are important for transcription, and its regulation by viral IE-gene products.

Previous work on HSV-1 Early gene regulation has concentrated on the response to superinfection of the viral thymidine kinase (tk) promoter in transformed cells, and the important sequence elements within this promoter for its constitutive function in Xenopus oocytes (see ref 12 for refs). Recently, a system to study the gD promoter by quantitative S1 mapping of RNA produced from plasmid constructions in a short-term transfection assay was described (12). This study showed that an HSV-1 Early promoter in a plasmid responded to trans activation by a superinfecting virus in the same way as its virally located counterpart. The gD promoter in a hybrid construction with the rabbit β-globin gene was activated during viral infection to produce RNA from the same cap-sites, and following the same time-course, as the normal gD gene. This viral regulation of transcription required only 83bp of HSV DNA upstream from the gD cap-sites for full activity, but deletion of these sequences led to progressive reductions in transcription (12). A similar sequence requirement has recently been demonstrated for viral activation of the tk promoter (13). Neither study could distinguish sequences which might be required specifically for regulation by Vmw175 from those sequences involved in more general aspects of promoter function.

In this paper a number of deletion and insertion mutants within the essential region of the gD promoter are described. The transcriptional activity of the mutants was studied after trans activation by viral products and also after cis activation using the SV40 enhancer. Thus it was possible to search for sequence elements involved solely and specifically in trans activation. Surprisingly, no specific sequence was detected. On the contrary, the integrity of the whole promoter region was necessary for full activation.

This study identifies several sequences that are involved in fully efficient transcription from the gD promoter. These include G-rich upstream signals which can function in both orientations, a classical TATA-box element (14), a sequence that is not homologous to the canonical TATA-box but which may have a

role in determining the position of some RNA 5' ends, a TATA homology that does not act as a TATA-box, and a complex cap-site region which plays some quantitative role in transcription and appears to include pyrimidine starts.

MATERIALS AND METHODS

1. Bacteria, Cells and Viruses.

These were as described previously (12).

2. Plasmids.

The following plasmids were described in ref.12 :pRED2, pRED111, pRED119, pRED121, pRED122 and pRED123. pβ(244+)β (15) was supplied by W.Schaffner.

3. Construction of Recombinant Plasmids.

Because this paper describes a large number of recombinant plasmids their construction will be described briefly and complete structures given only in the most important cases. Full details are available from the author on request.

(i) 5' Deletion Mutants. The general structure of the pRED, pDER and pERD plasmids is shown in Figure 1. The construction of the 5' deletion mutants in the gD promoter region of pRED2 has been described (12); pRED120 is another plasmid in this series with an end-point at position +5 as defined in Figure 2.

(ii) 3' Deletion Mutants. pRED21 is a derivative of pRED2 containing a XhoI linker at position +16. The pDER series of 3' deletions was made by cutting pRED21 with XhoI, treating with Exonuclease III then Mung Bean nuclease and ligation with XhoI linkers as described (12). The deletion end-points were initially located by restriction enzyme analysis and then in all cases by DNA sequencing (16).

(iii) Combination Deletions; the pERD series. The pRED series 5' deletions retain the complete β-globin region of pRED2 (figure 1) and varying amounts of the gD promoter, but have lost all HSV sequences upstream of the XhoI linker which marks the deletion end-point. Conversely, the pDER mutants retain the upstream region but have lost part of the globin gene and downstream portions of the gD promoter. Because a XhoI linker was also incorporated at the deletion in these constructions (creating a unique XhoI site in both instances) the pRED and pDER deletions could be combined to give small, internal deletions within the gD promoter region. In each case, pREDx and pDERy were cut with XhoI and PstI (also a single cut enzyme); ligation resulted in pERDy.x. These plasmids contain identical HSV sequences (excluding the deletion) between positions -2045 and position +17 (including the homology at the HindIII site) of the gD promoter region (figures 1 and 2).

(iv) Combination Deletions Containing the SV40 Enhancer. pSVD1 was constructed in an analagous fashion to pSVD2 (12) except that pMD102 (17) was used as vector (see (12) for details). pSVD31 is a derivative of pSVD1 in which the BamHI site at the junction of the SV40 and HSV sequences has been converted to a SstI site by linker insertion (figure 1). pSVD5 was made from pSVD1 by deletion of the BamHI-SstI fragment (equivalent to that in pSVD2; see figure 1 ref. 12) and insertion of a XhoI linker at this position. Thus the ClaI-XhoI fragment of pSVD5 containing the SV40 enhancer could be exchanged with that of

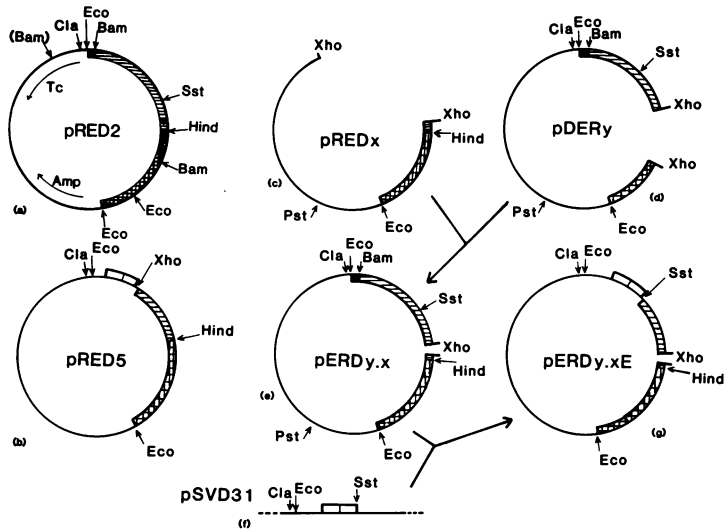


FIGURE 1. Structures of the principal plasmids used. (a): pRED2 with relevant restriction sites shown. The vector moiety (thin line) is pRE3 which has lost the BamHI site (in brackets)(20), the single- and double-hatched regions are HSV gD and rabbit β-globin DNA respectively, and the filled area is SV40 DNA (12). (b): pRED5; the SstI-EcoRI fragment of pRED2 has been replaced by SV40 DNA including the 72bp repeat enhancer (open boxes) with a XhoI site created at the junction. (c): pRED series of 5' deletions retaining variable proximal portions of the gD promoter. (d): pDER series of 3' deletions containing variable distal parts of the gD promoter. (e): pERD series combination deletions made from pRED and pDER plasmids and containing a small deletion in the gD promoter region represented by a gap. (f): The ClaI-SstI region of pSVD31 containing the SV40 enhancer (open boxes). (g): The pERD.E series plasmids which contain the SV40 enhancer fragment from pSVD31 and the same deletion as the corresponding pERD plasmid.

pRED4 (12) to give pRED5 (figure 1). pRED5 was used as the wild-type control for the pERD.E plasmids; all have the enhancer at exactly the same distance from the promoter region. To insert the enhancer in the pERD plasmids, the ClaI-SstI fragment of pERDy.x was replaced with that from pSVD31 (containing the enhancer) to give pERDy.xE (figures 1 and 2).

(v) Inversion of the Upstream region of the gD Promoter; pRED250. pRED7.122 was cut by XhoI and SstI and ligated under cohesive end conditions to the XhoI-AhaIII fragment of pRED119 containing bases -23 to -83. Then the 3' protusions of the SstI site were removed by the 3'-5' exonuclease of DNA Polymerase I Klenow fragment so that this could be blunt-end ligated to the AhaIII site, in the presence of SstI linkers, to regenerate the SstI site and re-circularise the molecule. The final structure

is shown in figure 2. This was confirmed by DNA sequencing. pRED250E was constructed by exchange of the ClaI-SstI fragments of pSVD31 and pRED250; in this case the SV40 enhancer is present immediately upstream of the gD promoter (position -100) instead of position -392 in the pERD.E plasmids.

(vi) pRED211 and pRED213; Linker Insertion Mutants of the TATA-box. pRED2 was cut partially with AhaIII (which cuts the sequence TTT/AAA) and full length linear molecules isolated from a sucrose density gradient. These were ligated with XhoI linkers. pRED211 was isolated as a full length plasmid containing a XhoI site instead of the AhaIII site which makes up part of the TATA-box at position -25. DNA sequence analysis showed that 5 tandem linker inserts were present; pRED213 was isolated after extensive cutting of pRED211 with XhoI and religation. DNA sequencing showed that a single XhoI linker was now present. pRED211E and pRED213E were made by exchange of ClaI-SstI fragments with pSVD31 as described above.

4. Calcium Phosphate Transfection, Infection of HeLa Cells, RNA Isolation and S1 Mapping Analysis.

The methods used were as described (12). In all experiments 10μg of the plasmid under test was co-transfected with 10μg of the internal control plasmid, pβ(244+)β . Viral infections were performed 24h. after transfection in trans activated experiments and RNA prepared 4h. after infection. In cis activated experiments with the pERD.E plasmids no virus was added, and RNA made 28 hours after transfection. The probe used for S1 mapping was a single stranded DNA fragment labelled at the BstNI site at position +136 in the globin gene of pRED2 (fig 2). This probe was used for experiments using plasmids whose mutations did not extend into the cap-site region. Because of the break in homology between the probe and RNA molecules initiated upstream of the mutation in the plasmid (see Discussion) a band appears after S1 analysis which maps to the site of the mutation (see figs. 3 and 4). Therefore, to eliminate this junction band from the cap-site region (which would otherwise invalidate the calculations), when recombinants containing mutations in the cap-site region were used, the probe was made from the same mutant plasmid. Suitably exposed autoradiographs were scanned with a densitometer and the areas of the peaks quantitated using a tablet digitiser. The ratio of correctly initiated 5' ends at the gD promoter compared with the globin 5' ends of the internal control (pβ(244+)β) was calculated. In trans activated experiments the value of this ratio obtained with pRED2 was taken as 100 and the ratios obtained with the mutants expressed as a percentage of this ratio. Similarly, in cis activated experiments, the ratio obtained with pRED5 was taken as 100. The activities given in figure 2 and Table 1 represent the averages obtained from at least 4 independant experiments using 2 different preparations of DNA.

RESULTS

Analysis by Quantitative S1 Mapping of Deletion and Insertion Mutants within the gD Promoter.

A preliminary analysis of the DNA sequence requirements

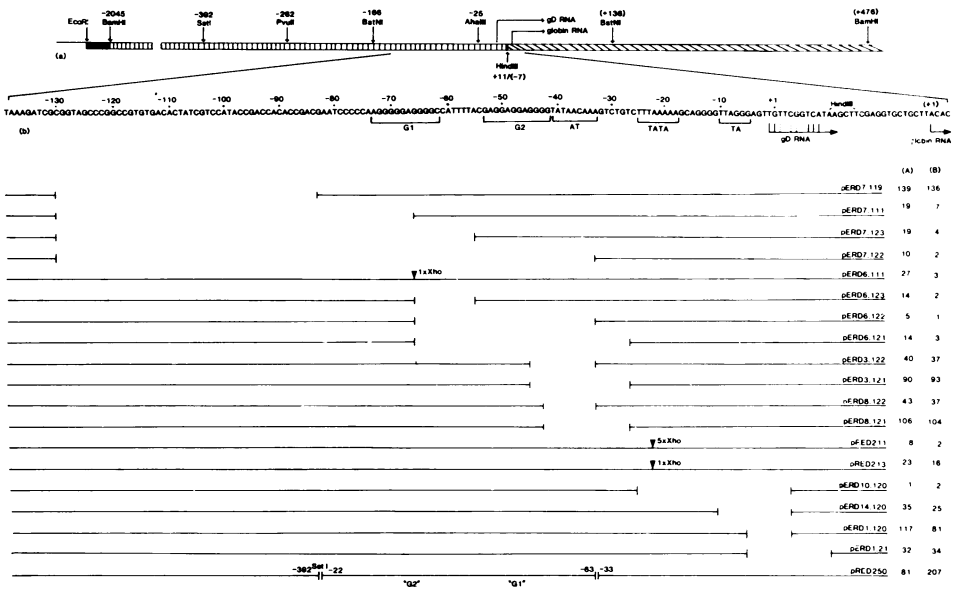


FIGURE 2. Detailed structures of pRED2 and the mutants used. (a): pRED2 with relevant restriction sites. The probe for S1 mapping was a single strand labelled at the BstNI site at (+136), isolated using the BamHI-PvuII fragment. Vertical hatched line indicates HSV sequences, diagonally hatched represents rabbit β-globin. The cap-site region of the gD promoter is shown, and that of the β-globin promoter is also indicated although it is only used in the internal control pβ(244+)β(12). Coordinates in brackets refer to globin sequences, the others to HSV DNA, both numbered in relation to their cap-sites. (b): DNA sequence of the gD promoter region showing the cap-sites in detail and other sequence features discussed in the text. The next 19 lines represent the details of the mutants. Deletions, which all have a XhoI linker (CCTCGAGG) inserted, are shown as a gap, insertions as a filled triangle. The mutants are identified on the right. In pRED250 the G-rich groups G1 and G2 are in opposite orientation to pRED2. (A) Relative transcriptional activities of the mutants after viral activation; pRED2=100. (B) Relative transcriptional activities of the pERD.E versions of the mutants when activated in cis with the SV40 enhancer; pRED5=100 (See Table 1).

for the virally regulated transcription from the gD promoter has been published (12). A short-term transfection assay was used to study the activation of the gD promoter (linked to the rabbit β-globin gene in a plasmid construction) by quantitation of correctly initiated RNA. The same assay procedure has been adopted to make a detailed analysis of the role of sequences

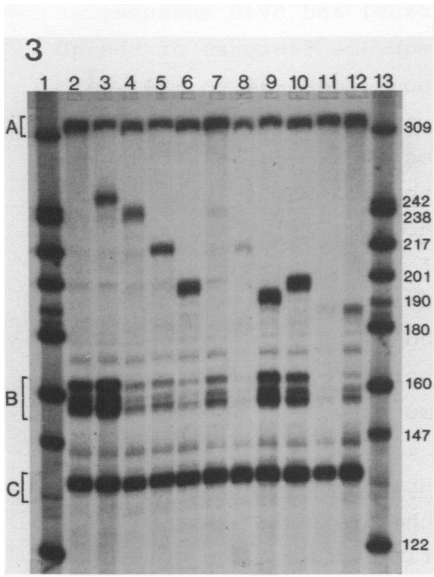
within the gD promoter in virally (trans) and SV40 enhancer (cis) activated transcription. The sequence features of the gD promoter include an AC-rich region (positions -94 to -84), a G-rich region (positions -63 to -73) and a closely related G-rich area (-42 to -53), TATA homologies at -33 to -41 and -18 to -25, and a third G-rich area (-3 to -17) before the cap-sites at positions +1 to +11 (figure 2). The results with deletion and insertion mutations affecting each of these sequences, and the intervening areas, will be considered in turn. In many constructs the number of bases deleted roughly corresponds to those replaced by the XhoI linker. Therefore some could be considered multiple clustered point-mutants, but are called deletions for ease of reference.

Nucleotides -130 to -83 are not Essential for Expression.

It has been shown that all the sequence elements necessary for fully activated transcription from the gD promoter lie wihin 83bp of the RNA cap-sites (12). This finding was supported using the pERD7 series plasmids, which all have deletions with upstream end-points at position -130 (figure 2). pERD7.119 has lost bases -129 to -84, which include the AC-rich region which has been found upstream of many HSV-1 Early promoters (18). Transcription from this promoter was unimpaired after viral activation and also when stimulated by the SV40 enhancer in the absence of any viral products (pERD7.119E) (figures 2, 3 (track 3) and 4 (track 3) and Table 1). In fact, the activities of both these plasmids were reproducibly above the control value. This small effect may indicate the presence of inhibitory sequences in this area.

The G-rich Sequence at -63 to -73 is an Important Element of the gD Promoter.

The deletion in pERD7.111 extends between bases -129 and -68 inclusive (figure 2). This removes the upstream part of the G-rich sequence (G1) and results in a significant drop in trans activated transcription (Table 1). In pERD7.123 the whole of this element is removed and a similar reduction was observed (figure 2). The pERD6 plasmids have their upstream end-point at position -68 (figure 2). Thus pERD6.111 has a XhoI linker in the middle of this run of G residues, and pERD6.123 has a XhoI



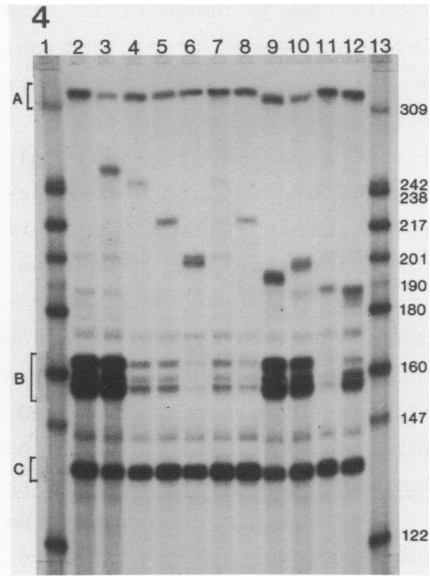


FIGURE 3. (left) Quantitative S1 mapping of trans-activated plasmids. Tracks 1 and 13; molecular weight markers. Tracks 2-12; pRED2, pERD7.119, pERD7.111, pERD7.123, pERD7.122, pERD6.111, pERD6.123 pERD8.121, pERD8.122, pRED211 and pRED213 respectively. A= band at full-length probe position, B= gD 5' ends, C= β-globin 5' ends.

FIGURE 4. (right) Quantitative S1 mapping of cis-activated plasmids. Tracks 1 and 13; molecular weight markers. Tracks 2-12; pRED5, pERD7.119E, pERD7.111E, pERD7.123E, pERD7.122E, pERD6.111E, pERD6.123E, pERD8.121E, pERD8.122E, pRED211E and pRED213E respectively. A= band at full length probe position, B= gD 5' ends, C= β-globin 5' ends.

linker replacing the downstream portion (figure 2). The results with these plasmids are consistent with those of the pERD7 series and confirm that the G-rich region G1 is an important element of the gD promoter. This region is not specific for viral activation because all these mutants were also reduced when activated in cis (pERD.E plasmids; figure 4 tracks 4,5 7,8 and Table 1). These relative reductions were larger than after viral activation (Table 1). This implies that both methods of promoter activation require this sequence but the details of the mechanisms may differ.

TABLE 1.

Plasmid	Mutation[a]	Activity[b]	Plasmid[c]	Activity[d]
pRED2	-	100	pRED5	100
pERD7.119	-129/-84	139	pERD7.119E	136
pERD7.111	-129/-68	19	pERD7.111E	7
pERD7.123	-129/-56	19	pERD7.123E	4
pERD7.122	-129/-34	10	pERD7.122E	2
pERD6.111	Xho -68/-67	27	pERD6.111E	3
pERD6.123	-67/-56	14	pERD6.123E	2
pERD6.122	-67/-34	5	pERD6.122E	1
pERD6.121	-67/-27	14	pERD6.121E	3
pERD3.122	-46/-34	37	PERD3.122E	40
pERD3.121	-46/-27	90	pERD3.121E	93
pERD8.122	-43/-34	43	pERD8.121E	37
pERD8.121	-43/-27	106	pERD8.121E	104
pRED211	5Xho -23/-22	8	pRED211E	2
pRED213	Xho -23/-22	23	pRED231E	16
pERD10.120	-25/+4	1	pERD10.120E	2
pERD14.120	-10/+4	35	pERD14.120E	25
pERD1.120	-4/+4	117	pERD1.120E	81
pERD1.21	-4/+11	32	pERD1.21E	34
pRED250	inversion	81	pRED250E	207

(a) coordinates of deleted bases (inclusive), and position and number of linker insertions given. See text for details of pRED250. (b) transcriptional activities (as defined in Materials and Methods) after viral activation relative to pRED2. The figures given are the averages of at least 4 different experiments. The individual values varied only a few percent from the mean in most cases. (c) Plasmids contain the SV40 enhancer at position -392 and the same mutation as in column 2. (d) transcriptional activities in the absence of viral infection relative to pRED5. The notes in (b) also apply here.

Sequences from -26 to -55 are Non-essential, but are Required for Fully Active Transcription.

This region contains a second G-rich area (G2) closely related to the one considered above and which is deleted in plasmids pERD7.122 and pERD6.122 (figure 2). Both show further reductions in both trans and cis activated transcription (figure 3 track 6, figure 4 track 6 and Table 1) This suggests that this area contains additional promoter elements, although the major effect is due to loss of the upstream G-rich tract. Note that pERD6.121, although having lost more DNA than pERD6.122, was activated in trans (and perhaps also in cis) more effectively (figure 2 and Table 1). This point is considered below. Plasmids pERD3.122 and pERD8.122 contain similar deletions covering most

or part of G2; both show a moderate drop in assays using both systems of promoter activation (figures 2,3 (track 10), 4 (track 10) and Table 1). Surprisingly, when the deletions were extended to position -26 (pERD3.121 and pERD8.121) full promoter activity was restored, again in both assays (figures 2,3 (track 9), 4 (track 9) and Table 1). This result implies that either the spacing of the sequences in this area is important (although in pERD8.122 this is maintained close to that of pRED2 because of the XhoI linker insert) or in some way the XhoI linker can reconstitute the required structures in pERD6.121, pERD3.121 and pERD8.121. Alternatively, the sequence between -34 and -26 may contain an element which has an inhibitory effect on transcription.

One feature of these results is that the distribution of cap-sites is exactly the same as in pRED2 and pRED5 (figure 3 tracks 2,9 and 10, and figure 4 tracks 2,9 and 10). This demonstrates that the TATA homology between -33 and -41 does not function as a TATA-box even though it has excellent homology with the concensus sequence (14)(but see below).

The AT-rich Sequence Between -18 and -25 Acts as a Typical TATA-box.

This region was considered likely to be a functional TATA-box because of its distance from the cap-sites. pRED211 contains 5 tandem linker insertions in the AhaIII site at position -23 (figure 2). This mutation resulted in a 10- to 20-fold decrease in correctly initiated RNA during viral activation (Table 1). However, the most upstream cap-sites were reduced 5-fold more than the lowest group (figure 3 track 11). This effect is most easily seen with pRED213, which has a single linker inserted and expresses at about 20% (figure 3, track 12). The number of linker inserts clearly affects the level of transcription. This could be due to a "distance effect" analagous to those observed in the tk (19) and SV40 Early (20) promoters where reductions in promoter activity were observed when the upstream elements were moved more than about 40bp from their normal positions near the TATA-box. Because these mutations clearly affect cap-site usage it can be concluded that the TTTAAAAA sequence acts as a TATA-box in the gD promoter.

Again, there was no evidence that this region was specifically involved in viral activation because the cis activated plasmids pRED211E and pRED213E gave results quantitatively and qualitatively almost identical to their trans activated counterparts (figure 4 tracks 11 and 12, and Table 1).

Deletions in the TATA-box and Cap-site Regions Affect the Quantity and Position of the RNA Starts and Reveal an Additional Sequence Which Appears to Affect the Position of the RNA 5' Ends.

The results with plasmids pERD10.120, pERD14.120, pERD1.120 and pERD1.21, and their pERD.E variants are discussed in this section. The deletion in pERD10.120 extends between bases -25 to +4, a region which includes the TATA homology considered above, part of the cap-site region and the G-rich region between (figure 2). In trans activation experiments, this recombinant showed a drastic reduction in transcription (Table 1), with some faint bands appearing 30bp downstream of the TATA homology at position -41 (not shown). Therefore this sequence can act as a TATA-box at very low efficiency. Similar results were obtained with pERD10.120E (figure 5A track 1). Therefore the region -25 to +4 contains sequences which are essential for both cis and trans promoter activation.

These important sequences were defined more precisely using smaller deletions. pERD1.21 (deletion -4 to +11; figure 2) gave a 3-fold reduced level of transcription with novel cap-sites mapping in two groups at closely analogous distances downstream of position -25 compared to those of pRED2 (figures 5A (track 3) and 6). Therefore the normal cap-site region (completely deleted in this plasmid) appears to have some quantitative role. The spread of 5' ends with pRED2 may reflect either the sequences around the cap-sites themselves, or they may be directed by sequences further upstream. The appearance of two groups of starts with pERD1.21 indicates that the latter is more likely. The same quantitative and qualitative effects were observed with pERD1.21E (Table 1). pERD14.120 (deletion -10 to +4; figure 2) gave similar reductions in both cis- and trans-activation as pERD1.21 (figure 5B track 3, Table 1) but in this case only one group of caps were found, mapping at a

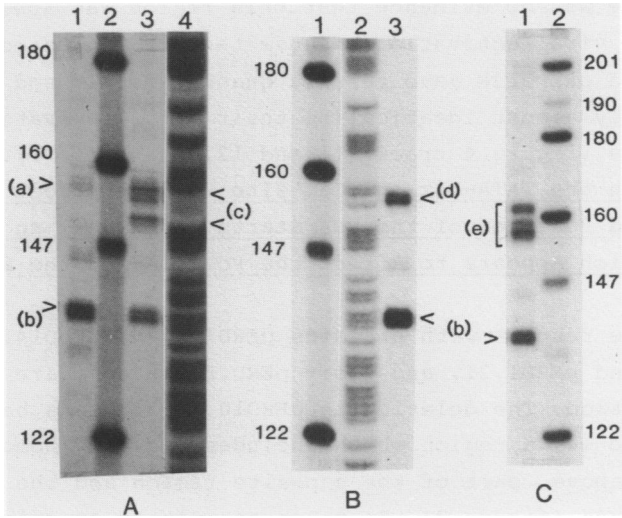


FIGURE 5. Quantitative S1 mapping of TATA-box and cap-site deletion plasmids. A. Track 1: pERD10.120E with 5' ends directed by the mutant promoter (a) and internal control (b) marked. Track 2: Molecular weight markers. Track 3: pERD1.21 5' ends marked (c). Track 4: A+G Sequence track of probe isolated from pERD1.21. B. Tracks 1-3: Molecular weight markers, A+G sequence track of probe derived from, and 5' ends (d) directed by pERD14.120 respectively. (b)= β-globin 5' ends. C. Track 1; 5' ends from pERD1.120 (e) and pβ(244+)β (b), Track 2; molecular weight markers.

position roughly corresponding to the upper group in pRED2 (figure 6). Therefore, in agreement with the results with pRED211 and pRED213 discussed above, the TTTAAAAA signal appears to be primarily involved in the initiation of the most upstream starts while the downstream bands occur in these experiments only when sequences between -10 and +4 (labelled TA in figure 2) are present. To confirm this observation, pERD1.120 was constructed (deletion -4 to +4; figure 2). This plasmid gave almost identical quantitative and qualitative results to pRED2 (figure 5C track 1, Table 1), as did pERD1.120E compared to pRED5 (not shown). In both cases the 5' ends were found at the same positions relative to the TTTAAAAA signal in pRED2. Therefore, the region -10 to -4 contains a sequence (TTAGGG) which, although not homologous to the canonical TATA-box sequence, appears to be involved in the positioning of some

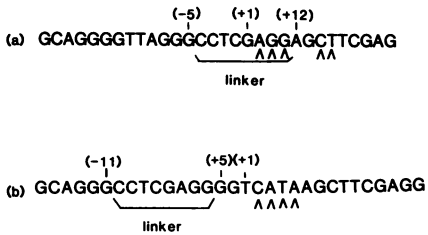


FIGURE 6. Positions of the major cap-sites of pERD1.21 and pERD14.120. (a) DNA sequence of probe derived from pERD1.21; (-5) and (+12) refer to the deletion end points with the XhoI linker between. (+1) indicates the relative position of the +1 coordinate in pRED2, counting the same distance downstream from the T at position -25. (b) As above, with pERD14.120. (-11) and (+5), deletion end points, (+1), relative +1 position in pRED2. Arrows indicate major 5' ends.

transcriptional initiation events. Its deletion in pERD14.120 results in the loss of a set of specific 5' ends, and a corresponding decrease in total transcription. Other examples of specific initiation of transcription directed by sequences not homologous to a TATA-box have been reported (21,22,23,).

However, if the cap-site region itself is primarily responsible for the pattern of 5' ends observed, an alternative explanation of these results is possible. The 8 nucleotides deleted in pERD1.120 are replaced by the Xho linker, thus maintaining the exact distance between the TATA-box and remaining cap-site region. Therefore the difference between pERD1.120 and pERD14.120 could be due to a difference in distance between the TATA-box and cap-site nucleotides remaining. In this analysis, the region labelled TA in figure 2 does not have a role in determining the position of the more downstream 5' ends. It is not possible to distinguish between these explanations on the available evidence, but it may be pertinent that only part of the cap-site region is present in pERD1.120. Therefore, although the RNA initiation pattern is the same as that of pRED2, the actual 5' ends used are different.

The Upstream Element of the gD Promoter is Equally Functional in Both Orientations.

An analysis of the DNA sequences upstream of a number of HSV-1 Early genes revealed few common homologies when the promoters were compared in the same orientation (12). However, when both orientations were considered, variants of the G-rich elements were present at various positions in all promoters.

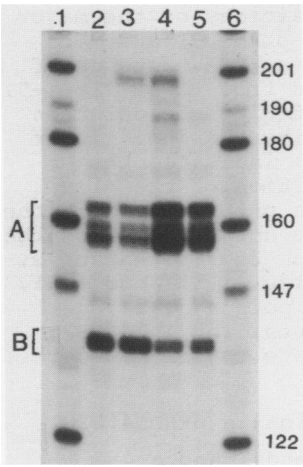


FIGURE 7. Quantitative S1 mapping of inversion mutants pRED250 and pRED250E. Tracks 1 and 6; molecular weight markers, Tracks 2-5; pRED2, pRED250, pRED250E and pRED5 respectively. A= gD 5' ends, B= β-globin 5' ends.

Since a C-rich sequence has been shown to be an essential element of the tk promoter (24), it seemed likely that these regions were functionally equivalent. To test this idea, pRED250 was constructed in which nucleotides -23 to -83 were inserted in reversed orientation between positions -392 and -33, thus maintaining the TATA-box but inverting the upstream region. When activated in trans, pRED250 responded almost as well as pRED2 (figure 7 and Table 1). The cis-activated inversion (pRED250E) was two-fold stronger than its control (pRED5), probably because the enhancer is next to the gD promoter instead of about 300bp upstream, as in pRED5 and the pERD.E plasmids. This is consistent with the documented "distance" effects with the SV40 enhancer (20).

DISCUSSION.

This paper describes experiments designed to elucidate the structure and regulation of the gD promoter. The basic features of the experimental system used have been described (12). The RNA produced from the hybrid gD/globin gene is produced from the same cap-sites and following the same kinetics as that from the gD promoter during normal viral infection. In the absence of an enhancer, transcription is entirely dependant on viral activation (12). Therefore the hybrid gene promoter is operating in a manner comparable to its viral counterpart. The

method of quantitation allows reliable comparison between different experiments. All transfections contained an internal control which was used to to standardise the activity of the test plasmid as a ratio; this ratio was compared with that of the wild-type plasmid to give as a percentage the relative transcriptional activity (see Materials and Methods, section 4). The transcriptional ratios observed with the same promoter mutants were different in trans and cis activated experiments (see figure 7). Therefore each series included its own "wild-type" plasmid for standardisation (pRED2 and pRED5). The two different series of plasmids were used because plasmids already activated in cis did not strongly respond to trans activation (results not shown). All transcripts mapping at the globin cap-sites were derived from the internal control plasmid (pβ(244+)β) while all those at the gD cap-sites came from the test plasmids (12). Other minor transcript bands were also apparent; these will be considered below. The ratio observed for pRED2/pβ(244+)β in many experiments varied only slightly (range 1.16-1.44; this variation was typical of all the data). Thus, although the globin promoter in pβ(244+)β was also activated in trans by HSV infection (12) this must be constant and therefore does not affect the quantitation. This analysis does not formally distinguish between activation and RNA stabilisation, but as similar trans activation by the Adenovirus E1A gene product has been shown to be due to an increased rate of transcription in isolated nuclei (24) this is likely to be the case here too.

The results described in this report have revealed a number of functional DNA sequences within the gD promoter, which in many respects are similar to those in other promoters. In particular, the G-rich sequence at position -70 of gD is an inverted variant of the C-rich motifs that are important for promoter function in many cases (for example 17,25-28 and refs therein). The gD promoter, like the tk promoter, appears to contain two upstream regulatory regions. However, the more downstream G-rich region is not so important as the other, despite their close homology (10/11 bases). This indicates that either their position or flanking sequences are also important.

This study also supports the notion that many upstream elements may be able to function in both orientations. Previously this had been shown only for the 21bp repeat region of the SV40 Early promoter (17), and the upstream region of the tk promoter (27). If this property of upstream regions is found to be widespread, functional bi-directionality can no longer be assigned solely to, and used as a defining feature of, enhancers. The clear difference between "upstream" regions and enhancers is that upstream regions need to be closely linked to a TATA-box for full activity (19,20) while enhancers do not. In addition, enhancers require upstream elements for full activity (20). The apparent functional bi-directionality of upstream regions, taken with the known latitude in exact distance of upstream elements from the TATA-box (19,20), may explain the seeming lack of general homology between upstream regions of the promoters so far analysed. Thus, direct or reversed homologies at various locations in upstream regions could have identical function. Indeed, C- or G-rich motifs within 100bp of the TATA-box are very frequent in promoter sequences. Therefore, sequences immediately upstream of the TATA-box that are important for full promoter activity may be quite well conserved. However, this hypothesis needs confirmation by further manipulation (paricularly inversion) of upstream sequences.

A striking feature of the gD promoter is the arrangement of the cap-sites in three groups covering about 10bp. These are primarily directed by the TTTAAAAA sequence at -25 as seen by the reduction in all 5' ends in mutants pRED211 and pRED213 (figure 3 tracks 11,12). Remarkably, the downstream bands do not appear when the TTAGGG sequence at -10 is deleted (pERD14.120; figure 5B), suggesting that this sequence has a role in positioning of some 5' ends. It is not clear why an excellent TATA homology at -41 is not involved in cap-site selection. The functional differences in these AT-rich regions must be controlled to some extent by their flanking sequences. Another unusual feature of the gD promoter is the decreased activity with a cap-site region deletion in both cis and trans activated experiments (pERD1.21 figure 2). This contrasts with the case of the tk and rabbit β-globin promoters where no

reductions were seen with analagous mutants (25,26). The gD promoter is particularly strong, producing much more RNA than the rabbit β-globin promoter when linked to the SV40 enhancer (figure 4 track 2). This increase is mostly due to an increased number of start-sites; it is possible that the nucleotide sequence around the gD caps allows particularly high rates of initiation so disruption of this arrangement leads to decreased promoter efficiency.

The aim of this investigation was to identify DNA sequences involved in trans-activation of transcription from HSV-1 Early promoters by viral IE-gene products. One possibility was that, as in some systems of trans-activation, regulation might be caused by recognition of a specific sequence within the promoter region (for example, Drosophila heat-shock (30) and hormone activated genes (31,32); see (33) for a review). These results strongly imply that this is not the case for the gD promoter. Presumably this is also true for other Early gene promoters. If there were a sequence which was solely concerned with trans activation (perhaps by direct interaction with viral IE polypeptides) it would have to be intermeshed with or overlapping the sequences which respond to cellular factors detected by the cis-activated plasmids described here. In view of the extensive mutational analysis conducted this is unlikely. The observations that the IE-gene products of HSV-1 (12), Pseudorabies Virus (34) and Adenovirus (35) can trans-activate a variety of unrelated promoters in transient expression assays (see below) support this conclusion. However, it is known that, apparently in contrast to host promoters, chromosomally located HSV Early promoters in transformed cell lines can be activated during virus infection (36,37,38). Using this technique a region thought to be specific for induction of HSV tk transcription has been proposed (39). However, different cell lines show wide variations in activation levels which would make certain identification of a sequence required specifically for induction difficult. Since these experiments have usually selected for activity of the tk promoter, it could be present in a local environment open for activation by Vmw175 even though it is in the chromosome.

The mechanism of trans activation by HSV-1 remains obscure. Recently it has been shown that the products of the Adenovirus IE gene E1A and a Pseudorabies virus (PRV) IE-gene also trans-activate not only the relevant Early genes but also a variety of unrelated promoters in plasmid constructions (24,34,35,40,). Since PRV and HSV are related Herpes viruses, and the HSV-1 IE-gene product Vmw175 is able to complement Adenovirus E1A mutants (41) it is probable that these examples of trans-activation share similar mechanisms. Possible explanations include (i) direct recognition of specific Early promoter sequences (ii) recognition of bifunctional or general promoter sequences, such as the TATA-box, leading to increased polymerase activity (iii) modification of the DNA template (iv) modification of RNA polymerase II (B) or other transcription factors (v) inactivation of host encoded inhibitors (42) (vi) transport of DNA that is not highly organised into chromatin to especially active regions of the nucleus, or (vii) release of transcription factors from host chromatin thus making them available for viral transcription. This study (given the proviso noted above) eliminates specific recognition of Early promoters because no class-specific sequence was detected. In addition, a completely unrelated promoter (rabbit β-globin) is also activated by Vmw175 ((12) and unpublished results). It is not possible to eliminate any of the other alternatives on the available evidence. However, a few common features of trans activation by these viruses provide some clues. The activation does not in general extend to chromosomal promoters (12,34,43) although E1A activates the host heat-shock genes (44). HSV-1 infection also activates the heat-shock response, but this is not thought to be mediated by Vmw175 (45). In addition to the promoter under study, the whole plasmid appears to become transcriptionally active. This can be seen in figures 3,4,5 and 7. Several reproducible bands are seen above the gD cap-sites in over-exposed autoradiographs and there is always a strong band at the position of the deletion; transcripts starting upstream of the deletion all map at this point because the probe was derived from an undeleted plasmid. Similarly, PRV activation resulted in transcripts starting at many locations, including

from vector sequences (34). This could be due to low frequency initiation directed by poor promoter homologues, which has been observed with the SV40 enhancer (46). Indeed, the pattern of extra bands is similar whether cis- or trans-activation was used (figure 7). The promoter sequence requirements for activation in cis or trans were similar, indicating that similar factors were required for transcription. However, trans-activation was always less affected by upstream mutations (Table 1)(34). This may indicate that some modification of RNA polymerase or other factors occurs during trans activation. These observations, together with the fact that the effects of cis and trans activation are not additive (results not shown) indicate that probably the same basic transcriptional machinery is used in both cases, but that it is activated in different ways. One intriguing possibility is that the IE proteins recognise DNA which is not highly organised into chromatin and either modify its structure so that it becomes readily accessible for transcription, or transport it to particularly active regions of the nucleus.

ACKNOWLEDGEMENTS. The author, a member of the Medical Research Council Virology Unit, thanks Professor J.H. Subak-Sharpe and Dr. J. Lindsay Whitton for helpful comments on the manuscript, Dr. W. Schaffner for pβ(244+)β, and Mr. Matt Dunlop for excellent technical assistance.

REFERENCES

1. Swanstrom,R.I. and Wagner,E.K. (1974) Virology 60, 522-533.
2. Clements,J.B., Watson,R.J. and Wilkie,N.M. (1977) Cell 12, 275-285.
3. Jones,P.C. and Roizman,B. (1979) J.Virol. 31, 299-314.
4. Lang,J.C., Spandidos,D.A. and Wilkie,N.M. (1983) EMBO J. in press.
5. Post,L.E. Mackem,S. and Roizman,B. (1981) Cell 24, 555-565.
6. Batterson,W. and Roizman,B. (1983) J.Virol. 46, 371-377.
7. Cordingly,M.G., Campbell,M.E.M. and Preston,C.M. (1983) Nucleic Acids Res. 11, 2347-2365.
8. Mackem,S. and Roizman,B. (1982c) J.Virol. 44, 939-949.
9. Preston,C.M. (1979) J.Virol. 29, 275-284.
10. Watson,R.J. and Clements,J.B. (1978) Virology 91, 364-379.
11. Watson,R.J. and Clements,J.M. (1980) Nature 285, 329-330.
12. Everett, R.D. (1983) Nucl. Acids Res. 11, 6647-6666.
13. Smiley,J.R., Swan,H., Pater,M.M., Pater, A. and Halpern,M.E. (1983) J.Virol. 47, 301-310.
14. Breachnach,R., and Chambon,P. (1981) Ann. Rev. Biochem. 50, 349-383.

15. de Villiers,J. and Schaffner,W. (1981) Nucleic Acids Res. 9, 6251-6264.
16. Maxam,A. and Gilbert,W. (1980) in Methods in Enzymology , Grossman,L. and Moldave,K. Eds. Vol.65, 43-62. Academic Press, New York.
17. Everett,R.D., Baty,D. and Chambon,P. (1983) Nucleic Acids Res. 11, 2447-2464.
18. Wagner, E.K. personal communication.
19. McKnight,S.L. (1982) Cell 31, 355-365.
20. Moreau,P., Hen,R., Wasylyk,B., Everett,R.D., Gaub,M.P. and Chambon,P. (1981) Nucleic Acids Res. 9, 6047-6068.
21. Piatek,M., Subramanian,K.N., Roy,P. and Weissman,S.M. (1981) J. Mol. Biol. 153, 589-618.
22. Cowie,A., Tyndall,C. and Kamen,R. (1981) Nucleic Acids Res. 9, 6305-6322.
23. Baker,C.C. and Ziff,E.B. (1981) J. Mol. Biol. 149, 189-221.
24. Leff,T., Elkaim,R., Goding,C.R., Jalinot,P., Sassone-Corsi,P., Perricaudet,M., Kedinger,C. and Chambon,P. personal communication.
25. McKnight,S.L. and Kingsbury,R. (1982) Science 217, 316-324.
26. Dierks,P., van Ooyen,A., Cochran,M., Dobkin,C., Reisser,J. and Weissman,C. (1983) Cell 32,695-706.
27. McKnight,S.L., Kingsbury,R.C., Spence,A. and Smith,M. in press.
28. Hen,R., Sassone-Corsi,P., Corden,J., Gaub,M-P. and Chambon,P (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 7132-7136.
29. Elkaim,R., Goding,C. and Kedinger,C. (1983) Nucl. Acids. Res. 11, 7105-7117.
30. Pelham,H.R.B. and Bienz,M. (1982) EMBO Journal 1, 1473-1477.
31. Dean,C. Knoll, B.J., Riser,M.E. and O'Malley,B.W. (1983) Nature 301, 551-554.
32. Majors,J. and Varmus,H.E. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 5866-5870.
33. Davidson,E.H., Jacobs,H.T. and Britten,R.J. (1983) Nature 301, 468-470.
34. Green,M.R., Treisman,R. and Maniatis,T. (1983) Cell 35, 137-148.
35. Imperiale,M.J., Feldman,L.T., and Nevins,J.R. (1983) Cell 35, 127-136.
36. Leiden,J.M., Buttyan,R. and Spear,P.G. (1976) J.Virol. 20, 413-424.
37. Kit,S. and Dubbs,D.L. (1977) Virology 76, 331-340.
38. Buttyan,R. and Spear,P.G. (1981) J.Virol. 37, 459-472.
39. Zipser,D., Lipsich,L. and Kwoh,J. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 6276-6280.
40. Weeks,D.L. and Jones,N.C. (1983) Mol. Cell. Biol. 3, 1222-1234.
41. Feldman,L.T., Imperiale,M.J. and Nevins,J.R. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 4952-4956.
42. Nevins,J.R. (1981) Cell 26, 213-220.
43. Babich,A., Feldman,L.T., Nevins,J.R., Darnell,J.E. and Weinberger,C. (1983) Mol. Cell. Biol. 3, 1212-1221,
44. Nevins,J.R. (1982) Cell 29, 913-919.
45. Notorianni,E.L. and Preston,C.M. (1982) Virology 123, 113-122.
46. Wasylyk,B., Wasylyk,C., Augereau,P. and Chambon,P. (1983) Cell 32, 503-514.