A detailed mutational analysis of Vmw110, a *trans*-acting transcriptional activator encoded by herpes simplex virus type 1

R.D. Everett

MRC Virology Unit, Church Street, Glasgow G11 5JR, UK

Communicated by J.H. Subak-Sharpe

Viral genes expressed during infection of tissue culture cells by herpes simplex virus type 1 (HSV-1) can be divided into three broad temporal groups called immediate–early (IE), early and late. Prior expression of IE gene products is required before later classes of genes can be transcribed. Using cloned copies of viral genes in plasmid vectors, it has been shown that the products of IE genes 1 and 3 (Vmw110 and Vmw175) can activate transcription from both viral and cellular promoters in short-term transfection assays. The effects of these two IE gene products are markedly synergistic in HeLa cells; the activation observed with both genes present is about 20 times that observed with Vmw175 alone. The mechanism of this activation and the properties of the IE proteins that are required are not well understood. This paper describes the construction of a large number of in-frame insertion and deletion mutations in a plasmid-encoded copy of IE gene 1. The ability of the mutant Vmw110 polypeptides to activate gene expression (in the presence of Vmw175) from the HSV-1 glycoprotein gD promoter linked to the chloramphenicol acetyl transferase gene was studied. The results show that the structural integrity of at least five regions of the polypeptide are important for its function in the presence of Vmw175.

Key words: herpes simplex virus type I/immediate–early genes/Vmw110 and Vmw175/transcription activation

Introduction

Herpes simplex virus type 1 (HSV-1) is a large enveloped virus containing a double-stranded DNA genome of ~150 kb which encodes ~75 genes (see Wagner, 1985, for a review). These genes can be divided into at least three classes depending on their temporally regulated expression. The five immediate–early (IE) genes are transcribed first and prior IE-gene expression is required before the later classes of genes, broadly described as early and late, can be transcribed (Honess and Roizman, 1974; Clements *et al.*, 1977; Jones and Roizman, 1979). Several of the IE genes are involved in the regulation of later classes of genes. The product of IE gene 3, Vmw175, is crucial for HSV-1 gene regulation since virus mutants with temperature-sensitive lesions in this gene fail to express early and late gene products at the non-permissive temperature (Preston, 1979) and revert to the pattern of gene expression seen at the IE stage of infection on shifting from permissive to non-permissive conditions (Watson and Clements, 1980). The products of IE genes 2 and 4 (Vmw63 and Vmw68) appear to be involved in the regulation of subsets of viral genes (Sears *et al.*, 1985; Sacks *et al.*, 1985; Everett, 1986). Recent studies using isolated IE genes in short-term transfection experiments have shown that the product of IE gene 1, Vmw110, is also able to activate HSV gene expres-

sion (O'Hare and Hayward, 1985a,b; Quinlan and Knipe, 1985; Gelman and Silverstein, 1985) and that its effect is considerably increased in the presence of Vmw175 (Everett, 1984, 1986). It has been suggested that Vmw110 is able to activate gene expression non-specifically (O'Hare and Hayward, 1985b), and it has been shown that the combination of Vmw175 and Vmw110 can activate a variety of cellular promoters (Everett, 1984, 1985).

IE gene 1 is encoded in the repeated sequences bounding the long unique region of the HSV-1 genome and is therefore diploid. Its product, Vmw110, is located in the nucleus, is phosphorylated and binds to DNA in crude cell extracts (Pereira *et al.*, 1977; Hay and Hay, 1980; Ackermann *et al.*, 1984). The role of Vmw110 during lytic infection of tissue culture cells is still not clear. Until recently there were no conditional mutants or deletions in this region although deletion mutants of both HSV-1 and HSV-2 containing only one copy of the IE gene 1 had been isolated (Poffenberger *et al.*, 1983; Harland and Brown, 1985). Recently HSV-1 viruses have been constructed in which the majority of the IE-1 coding region has been deleted in both copies of the gene (Stow and Stow, 1986; Sacks and Schaffer, 1987). Although a plasmid-encoded derivative of the IE gene 1 containing a large deletion had no detectable activity above background in a short-term transfection assay (Perry *et al.*, 1986), the virus containing this deletion was able to grow in tissue culture (Stow and Stow, 1986). The mutant plaqued relatively inefficiently compared with wild-type, particularly in some cell lines, and it showed a very high particle to pfu ratio; it has been suggested that the mutant exhibits a defect at a low multiplicity of infection which can be overcome by increasing the multiplicity (Stow and Stow, 1986). Therefore it appears that Vmw110 is not absolutely essential for growth in tissue culture, but that it plays an auxiliary role such that it is required for fully efficient infection.

The availability of a transfection system for the assay of IE-gene function allows a rapid and straightforward approach to the analysis of the structural domains of a protein that are important for its activity. This paper describes a detailed mutational analysis of Vmw110. The complete DNA sequence and mapping of the RNA (including the two introns) of the IE gene 1 have been described (Perry *et al.*, 1986). This information has been used as a basis of experiments to introduce a large number of in-frame deletion and insertion mutations into the coding region of a plasmid encoded IE gene 1. The activities of the altered polypeptides were then studied in short-term transfection assays in the presence of Vmw175. The results have shown that the structural integrity of at least five regions of the polypeptide are important for its function.

Results

Isolation of insertion mutations within the Vmw110 coding region

The coding region of IE gene 1 starts at the ATG at position 1864 in the DNA sequence of the long repeat section of the HSV-1 genome, and continues to the TAA at position 5090, excluding the introns between nucleotides 1921–2685 and

Fig. 1. A map of the HSV-1 genome and the positions of IE genes 1 and 3. (a) The HSV-1 genome (in the prototype orientation) showing the long unique (U_L) and short unique (U_S) sequences bounded by internal (IR) and terminal (TR) repeat sequences. The positions and directions of transcription of the diploid copies of IE genes 1 and 3 are marked in the repeated sequences. The *HpaI* and *SstI* restriction sites are those bounding the IE gene 1 transcription unit shown in (b). (b) A selective restriction map of the *SstI*-*HpaI* fragment containing IE gene 1. The positions of the 5' and 3' ends of the IE-1 RNA, the predicted sites of initiation and termination of protein synthesis and the introns within the gene are marked. The nucleotide coordinates refer to those of the R_L sequence (Perry *et al.*, 1986).

Fig. 2. Construction of insertion mutations within the IE gene 1. Plasmid p111, containing the *SstI*-*HpaI* fragment including the IE gene 1 transcription unit cloned into pUC9, was partially cut with a restriction enzyme in the presence of ethidium bromide. Singly cut linear molecules were purified from a sucrose gradient and ligated to 12-bp *EcoRI* linker oligonucleotides. Plasmids with linker insertions were treated with *EcoRI* and re-ligated to reduce the insertion to a single linker. The identity of all the plasmids in the p110E series was confirmed by DNA sequence analysis.

Fig. 3. The positions of the insertion and deletion mutations used in this study. The top line indicates the nucleotide coordinates of the Vmw110 coding region and introns. The position of each insertion mutant is indicated by a vertical line and deletions are shown by a bar. Selected features of the predicted amino acid sequence are shown next to a scale indicating the amino acid number within the predicted sequence.

region. This was achieved by partial restriction enzyme digestion of p111 in the presence of ethidium bromide, isolation of singly cut linear molecules on a sucrose density gradient and ligation with a linker oligonucleotide. When the plasmid was cut by a restriction enzyme which produced a blunt-ended cut, a 12-bp *EcoRI* linker was inserted while when a 4 base 5' extension was produced, the ends were 'filled-in' with Klenow polymerase before ligation to an 8-bp *HindIII* linker. In either case, after elimination of multiple linker inserts, there is a net addition of 12 bp resulting in the insertion of four amino acids into the polypeptide. This scheme is outlined in Figure 2. In this study, enzymes *HaeIII*, *RsaI* and *FspI* were used to produce blunt-ended cuts while *AvaI* produced 4 base 5' extensions. It should be noted that this method can be modified for use with enzymes which produce other types of cuts simply by 'filling-in' or trimming back the unpaired bases and using a suitable linker to achieve an insertion of a number of bases which is divisible by three. This approach is essentially similar to that described as 'TAB' mutagenesis (Barany, 1985). A summary of the insertion mutants obtained showing their positions, insertions and modifications to the amino acid sequence is given in Figure 3 and Table I.

Isolation of in-frame deletion mutations within the IE gene 1 coding region

Small in-frame deletion mutations in the Vmw110 sequence in p111 were obtained in two ways. Firstly, adventitious deletion mutants were sometimes isolated during the screening of plasmids derived from an experiment designed to obtain insertion mutations. Examples of these were obtained after both *HaeIII* and *AvaI* mutagenesis. Secondly, deletion mutations could be constructed, using the unique restriction sites introduced by insertion mutagenesis, by joining the region 5' to one insertion to that 3' to a second insertion. Deletion mutants of the first group were sometimes out of frame when first isolated, but by deletion or addition of a further 4 bp (by filling-in or trimming the restriction site within the inserted oligonucleotide) it was always possible to obtain an in-frame derivative. The positions and deleted regions of the in-frame deletion mutants used in this study are summarized in Table II, and are shown diagrammatically in comparison with the positions of the insertions in Figure 3.

The assay - Vmw110 stimulates expression from the HSV-1 glycoprotein gD promoter synergistically with Vmw175

It has been shown that the cloned IE-1 and IE-3 gene products from HSV-1 Glasgow strain 17 syn⁺ stimulate the HSV-1 gD

Table I. A summary of the positions and activities of the insertion mutants

Mutant	Site ^a	Insertion site ^b and inserted oligo ^c	Insertion AA ^d and inserted AA seq. ^c	Activity ^f	Activity with multiple linkers ^g
p110R1	1886	<i>RsaI</i> (E)	8/9 (PEFG)	112 ± 14	
p110E25	1903	<i>HaeIII</i> (E)	14 (GPNSG)	100 ± 18	
p110E3-1	2855	<i>HaeIII</i> (E)	76/77 (PEFG)	87 ± 11	71 ± 19 (2)
p110A2	2939	<i>AvaI</i> (H)	105/106 (STSL)	64 ± 13	
p110F1	3074	<i>FspI</i> (E)	150 (PRIRG)	44 ± 2	
p110R2	3111	<i>RsaI</i> (E)	162 (SRIRA)	39 ± 8	
p110E8-1	3189	<i>HaeIII</i> (E)	188 (ARIRA)	22 ± 5	11 ± 2 (4)
P110E13	3216	<i>HaeIII</i> (E)	197 (ARIRA)	38 ± 10	
p110R3	3261	<i>RsaI</i> (E)	212 (SRIRA)	50 ± 4	
p110E32-1	3291	<i>HaeIII</i> (E)	222 (ARIRA)	22 ± 4	7 ± 1 (3)
p110R4	3497	<i>RsaI</i> (E)	245 (VPNSG)	78 ± 13	
p110E31-1	3596	<i>HaeIII</i> (E)	278 (ARIRA)	36 ± 5	10 ± 4 (4)
p110E14-1	3633	<i>HaeIII</i> (E)	290/291 (PEFG)	97 ± 9	63 ± 9 (2)
p110E53-1	3728	<i>HaeIII</i> (E)	322 (GPNSG)	67 ± 7	42 ± 10 (4)
p110E15-1	3783	<i>HaeIII</i> (E)	340/341 (PEFG)	83 ± 23	75 ± 7 (2)
p110E29-1	3791	<i>HaeIII</i> (E)	343 (GPNSG)	96 ± 11	100 ± 39 (4)
p110E33	3871	<i>HaeIII</i> (E)	370 (ARIRA)	96 ± 20	
p110E51-1	3885	<i>HaeIII</i> (E)	374/375 (PEFG)	94 ± 16	68 ± 15 (4)
p110A3	3923	<i>AvaI</i> (H)	388 (GRQAW)	72 ± 14	
p110E38-1	3943	<i>HaeIII</i> (E)	394 (ARIRA)	61 ± 11	46 ± 3 (4)
p110E28-1	3979	<i>HaeIII</i> (E)	406 (ARIRA)	15 ± 3	11 ± 6 (4)
p110A4	4031	<i>AvaI</i> (H)	424 (GRQAW)	51 ± 13	
p110E12-1	4045	<i>HaeIII</i> (E)	428 (ARIRA)	58 ± 3	43 ± 2 (4)
p110A5	4156	<i>AvaI</i> (H)	465 (VGKLG)	75 ± 7	
p110E30	4164	<i>HaeIII</i> (E)	466/467 (PEFG)	77 ± 13	
p110E45-1	4188	<i>HaeIII</i> (E)	474/475 (PEFG)	75 ± 12	134 ± 43 (2)
p110E56	4289	<i>HaeIII</i> (E)	509 (GPNSG)	64 ± 4	
p110E1-1	4404	<i>HaeIII</i> (E)	548/549 (PEFG)	103 ± 11	74 ± 3 (4)
p110A6	4406	<i>AvaI</i> (H)	549 (GRQAW)	95 ± 7	
p110E6-1	4486	<i>HaeIII</i> (E)	575 (ARIRA)	57 ± 9	11 ± 1 (3)
p110E17-1	4507	<i>HaeIII</i> (E)	582 (ARIRA)	45 ± 8	24 ± 4 (5)
p110A7	4534	<i>AvaI</i> (H)	592 (VGKLG)	50 ± 12	
p110E52-1	4543	<i>HaeIII</i> (E)	594 (ARIRA)	70 ± 12	14 ± 4 (3)
p110E4-1	4605	<i>HaeIII</i> (E)	614/615 (PEFG)	85 ± 15	118 ± 29 (4)
p110E9	4614	<i>HaeIII</i> (E)	617/618 (PEFG)	110 ± 19	
p110E58-1	4660	<i>HaeIII</i> (E)	633 (ARIRA)	32 ± 8	19 ± 3 (3)
p110E23	4676	<i>HaeIII</i> (E)	638 (GPNSG)	20 ± 3	
p110A8	4699	<i>AvaI</i> (H)	647 (VGKLG)	26 ± 5	
p110R5	4801	<i>RsaI</i> (E)	680 (SRIRA)	18 ± 2	
p110E35	4849	<i>HaeIII</i> (E)	696 (ARIRA)	30 ± 6	
p110A9	4916	<i>AvaI</i> (H)	719 (DRQAW)	18 ± 3	
p110R6	4921	<i>RsaI</i> (E)	720 (SRIRA)	15 ± 1	
p110E2	4930	<i>HaeIII</i> (E)	723 (ARIRA)	14 ± 2	

^aThe coordinate given is the position of the first base in the recognition sequence of the insertion site in the DNA sequence of IR_s as described by Perry *et al.* (1986).

^bThe restriction enzyme site at which the insertion occurs.

^cThe inserted oligonucleotide was either the 12-mer *EcoRI* linker 5'-CCCGAATTCGGG-3' (E) or the 8-mer *HindIII* linker 5'-CAAGCTTG-3' (H).

^dThe insertions lie either within the codon of the indicated amino acid or between the two codons shown.

^eWhen the insertion has disrupted a codon, that amino acid is deleted and replaced by the five amino acids shown. Insertion between codons simply adds four amino acids. The single letter amino acid code has been used.

^fThe activation of the gD promoter (using pgDCAT in conjunction with Vmw175 provided by p175) is given as a percentage of that obtained in parallel experiments with p111. The mean of at least four independent determinations is given, with the standard error of the mean. These data are plotted on Figure 5. The background level of activation obtained with p175 alone (6.7% ± 0.9) has not been subtracted from these values.

^gThe activities of insertions with multiple linker inserts are shown with the standard error of the mean. The figure in brackets refers to the number of linkers inserted as determined by restriction enzyme digests.

promoter synergistically in short-term transfection experiments using HeLa cells (Everett, 1984). This assay has been used to examine the phenotype of the insertion and deletion mutants in Vmw110. In these experiments, the gD promoter was linked to the chloramphenicol acetyl transferase gene in plasmid pgDCAT (Everett, 1986) and Vmw175 was provided by plasmid p175, which contains the IE gene 3 coding region linked to the SV40

early promoter and enhancer (Figure 4). This particular plasmid was chosen to provide Vmw175 because the SV40 early promoter has been shown to respond poorly to activation by HSV-1 IE gene products (O'Hare and Hayward, 1985a) and therefore this construction minimizes any problems in the interpretation of the results which might arise from the regulation of the IE 3 promoter by Vmw175 and Vmw110 (O'Hare and Hayward,

Fig. 4. Details of the CAT assay to measure the activation of gene expression. (a) Plasmid p175 contains the IE gene 3 coding region linked (at the *Bam*HI site at position +27 relative to the IE 3 RNA start sites) to the SV40 early promoter and enhancer region. (b) Plasmid pgDCAT contains the gD promoter linked to the CAT coding region and a polyadenylation signal. (c) Transfection of HeLa cells with pgDCAT alone (track 1), pgDCAT with either p175 or p111 respectively (tracks 2 and 3) and pgDCAT, p175 and p111 (track 4).

1986b). Typical results using this assay are shown in Figure 4c. When HeLa cells were transfected with pgDCAT alone, the resultant cell extracts contained only low levels of CAT activity (track 1). This low level of activity was increased by ~3- to 5-fold when either p175 or p111 was included in the transfection with pgDCAT (tracks 2 and 3). However, when both p111 and p175 were present in the transfection, the activity obtained from pgDCAT was on average about 20 times that obtained with activation by p175 alone (track 4). This 20-fold increase was used as the assay to test the phenotypes of the insertion and deletion mutants in the Vmw110-coding region of p111.

Analysis of the insertion mutants in Vmw110

The 43 insertion mutants in Vmw110 shown in Table I and Figure 3 were tested for their ability to activate the gD promoter in the presence of Vmw175. At least two independent preparations of each plasmid DNA were used for a total of at least four transfection experiments. In each transfection experiment, two positive control plates with p111 were prepared and the average activity (percentage chloramphenicol acetylated per mg protein extract in a 30-min incubation) from these two plates was used to calculate the relative activities of the mutants. The results are also given in Table I, together with the standard errors of the means of the individual values, and the data are plotted graphically in Figure 5. Activation of pgDCAT by p175 alone gave on average an activity of $6.7\% \pm 0.9$ of that with both p175 and p111 (average of 26 determinations); this background level has not been subtracted from the figures shown.

The first point to note about these results is that 27 of the 43 insertion mutants have 50% or more of the activity of p111. Therefore the effect of the insertion of four amino acids, often accompanied by the replacement of another (Table I), is frequently tolerated by the polypeptide and does not cause a substantial fall in activity. Secondly, the mutations frequently exhibit activities similar to those of their immediate neighbours, such that the polypeptide can be divided into regions in which insertions either do or do not affect function. These regions have been labelled

Fig. 5. (a) The relative activities of the insertion mutants within the Vmw110 coding region, compared with the wild-type gene, are plotted against their positions in the predicted amino acid sequence. The two open arrow heads show the positions of the two introns. The dotted line indicates the background level of activity obtained with pgDCAT and p175 in the absence of plasmids expressing Vmw110. The regions marked 1–5 at the top of the figure refer to the regions where insertion results in loss of activity as discussed in the text. (b) The activities of the multiple linker insertion plasmids are shown on the same scale with the same annotations. The numbers below the bar refer to the number of the p110E plasmid whose activity is represented by the line above the bar; since all three parts of this figure are printed exactly in line, these numbers can be used to identify to which plasmids the lines in part (a) refer. The arrow at the top of the line of p110E45 indicates that this value is above the scale (see Table I). (c) The activities of the deletion mutations shown to the same scale. The length of the bar indicates the extent of the deletion and its level the activity of the plasmid.

Table II. The positions and activities of the deletion mutants

Mutant	Deleted bases ^a	Deleted amino acids ^b	Inserted oligo ^c	Activity ^d
p110de11	2871–3224	82–199	–	18 ± 4.5
p110D13/32	3218–3291	197–222	<i>EcoRI</i>	22 ± 3.7
p110E18	3945–4046	394–428	<i>EcoRI</i> (×3)	7 ± 1.4
p110E50F	3929–3980	389–406	<i>EcoRI</i> + 4 bp	22 ± 4.7
p110E37	4047–4103	428–447	<i>EcoRI</i>	86 ± 5.6
p110D48/1	4382–4405	541–548	<i>EcoRI</i>	66 ± 1.1
p110D35/2	4851–4931	696–723	<i>EcoRI</i>	21 ± 3.8

^aCoordinates of the deleted bases inclusive.^bNumbers of the deleted amino acids inclusive.^cMost deletions contain a single 12-bp oligonucleotide inserted at the position of the deletion; that in p110E50F has an additional 4 bp created by filling in the *EcoRI* site.^dRelative activity compared to p111 in transfections containing pgDCAT and p175. The activation obtained with p175 by itself (6.7% ± 0.9) has not been subtracted from these values.

ed 1–5 in Figure 5a. The clearest example of a region in which all insertions lead to a significant loss of activity is region 5 towards the carboxy terminus of the protein; all insertions downstream of amino acid 633 show significantly reduced levels of activity. (For the purposes of this analysis, a significant reduction is taken as a figure of 50% or less, with greater weight given to those regions which show reductions with two or more contiguous mutations.) A second region in which insertions reduce activity is in the second exon, between amino acids 150 and 222 (region 1). In addition, there are three other regions where insertion also leads to a significant fall in activity. These are around position 278 (region 2) and between positions 394 and 424 (region 3), and 575 and 592 (region 4). These points will be discussed in more detail below. Finally, although this set of mutants is fairly evenly spread over the whole of the polypeptide, it should be noted that other regions that are important for function may exist in the gaps that remain.

The effect of insertion of multiple linker oligonucleotides

In some cases (all p110E plasmids with numbers of the form n-1) the plasmids initially isolated after linker mutagenesis contained multiple (two to five copies) linker inserts at the site of mutation. A study of the activity of these plasmids in cotransfections with pgDCAT and p175 indicated the following (see Table I and Figure 5b): (i) multiple linker insertions in regions of the protein where single insertions had only small effects on function resulted in no or only small further reductions in activity (plasmids p110E3-1, 15-1, 29-1, 51-1, 45-1, 1-1 and 4-1); (ii) multiple insertions at positions in the second exon and carboxy terminal region, in which single linker insertions caused significant reductions in activity, resulted in further (~2-fold) reductions in activity (plasmids p110E8-1 and 32-1 and 58-1); (iii) multiple insertion at the position of the insertion of p110E31-1 reduced activity to only 10%, while multiple insertion at the nearby sites of p110E14-1 and 53-1 resulted in activities of <50% – these data confirm that region 2 of the polypeptide is important for its function; (iv) similarly, multiple insertion at the site of p110E28-1 and the nearby sites of p110E38-1 and 12-1 also resulted in further reductions in activity (region 3); and (v) multiple insertions in the serine-rich region (see below; plasmids p110E6-1, 17-1 and 52-1; region 4) were all considerably less active than their single insertion variants (11, 24 and 14% respectively) – these data reveal the potential importance of this region more clearly.

In summary, the results obtained with the multiple linker insertion plasmids confirm and extend the conclusions drawn from the results with the single linker insertions; the polypeptide contains at least five regions in which disruption of the amino acid sequence results in significant loss of activity.

Assay of the deletion mutants in Vmw110

A small number of deletion mutants were constructed which complement the studies with the insertion mutants. The location of these deletions are shown in Figure 3, and their activities are given in Table II and Figure 5c. Deletion mutant p110de11 has been described before (Perry *et al.*, 1986) and has lost amino acids 82–199 from exon 2. This mutant has only a low level of activity in the presence of Vmw175, which is consistent with the insertion study since it has lost part of a region in which insertions reduce function (region 1). Interestingly, this deletion also removes the whole of a cysteine-rich region which has some homology both to a Varicella Zoster virus polypeptide and also to a 'metal finger' binding domain (Figure 3; Perry *et al.*, 1986; Berg, 1986). Deletion mutants p110D13/32 and p110D35/2 were both constructed in regions in which insertion resulted in loss of activity (regions 1 and 5 respectively), and as expected both deletions caused similar reductions. In contrast, deletions p110E37 and p110D48/1 occur in regions not markedly affected by insertion and neither shows a substantial loss of activity. Plasmid p110E28-1 gave only low levels of activation despite being surrounded by comparatively active insertion mutations; however, deletion mutants p110E18 and p110E50F have both lost the region of the insertion of p110E28-1 (region 3) and both show low levels of activity. Again, the results of the deletion and insertion studies are consistent and in this instance imply that there is a comparatively small region of the polypeptide around position 406 which plays an important role in protein function.

Visualization of the wild-type and mutant Vmw110 polypeptides in the transfected cell

One problem of probing for functional regions of a polypeptide by transfection of mutated plasmid genes is that it is not easy to be sure that no other unselected and undetected change has occurred in the plasmid, such that the polypeptide is either not made at all or is present in a truncated form. To overcome this problem, indirect immunofluorescence studies were performed on cells transfected with wild-type and mutant plasmids. A rabbit antiserum directed against a synthetic oligopeptide corresponding to the carboxy terminus of the Vmw110 polypeptide, and which is known to recognize the intact protein (Perry *et al.*, 1986) was incubated with formaldehyde-fixed, transfected HeLa cells. Fluorescein-conjugated sheep anti-rabbit Ig was then used to visualize the bound rabbit antibodies. Compared with a mock-transfected control (Figure 6a) plates that had been transfected with p111 showed a proportion of nuclei containing a number of brightly fluorescing spots (Figure 6b and c). Not all of the cells contained these spots, which is not surprising since transfection is never 100% efficient, and the spots were never seen in negative controls. This result confirms previous studies that have shown that Vmw110 is located in the nucleus (Pereira *et al.*, 1977; Ackermann *et al.*, 1984). Cells were transfected with each of the 50 deletion and insertion mutants described above. The results were similar in all cases; all gave rise to positive fluorescent nuclei that were similar in appearance to those obtained with p111. The conclusion of this experiment is that all the deletion and insertion mutants express a protein which contains the carboxy terminus of Vmw110, and therefore the plasmids are unlikely to contain other undetected lesions (such as a hidden frame

Fig. 6. Immunofluorescence of cells transfected with p111 and stained with a rabbit anti-Vmw110 antiserum. (a) Negative control showing mock-transfected HeLa cells. Occasional rounded cells show some diffuse fluorescence. (b) Cells transfected with p111. Several cells show distinct punctate nuclear fluorescence, particularly at the top of the figure. (c) An enlargement of some positively staining cells.

shift). In addition, since all of the mutant polypeptides were located in the nucleus, it follows that none of the mutations affect the cellular localization of the polypeptide. It is also of interest to note that neither the strength nor quality of the fluorescence was affected by the inclusion of p175 in these experiments (data not shown).

Assay of the effect of mutations in Vmw110 on the level of cytoplasmic RNA from the gD promoter

Assay of the effect of mutations in Vmw110 using the CAT system provides only an indirect measure of the level of RNA transcribed from the test gene. In order to show that the results of the CAT assays reflected the amount of RNA produced from the gD promoter and found in the cytoplasm, a series of transfections were performed using the plasmid pRED4, which contains the gD promoter linked to the rabbit β -globin gene (Everett, 1983). The gD promoter in pRED4 was activated by cotransfection of p175 with p111 or a selection of the insertion and deletion mutants of Vmw110. Total cytoplasmic RNA was prepared 2 days later and the amount of RNA initiated at the gD promoter quantitated by hybridization to a single-stranded DNA probe (prepared from pRED4) 5' end labelled at position +136 in the globin coding region, followed by digestion with S1 nuclease. The gD promoter is fused at position +11 to the globin gene at position -7 so, allowing for intervening linker DNA, RNAs initiated at the gD promoter should give protected hybrids of ~160 nucleotides.

The results are shown in Figure 7. The gD promoter did not give any detectable correctly initiated RNA in the absence of activation (track 2). Activation by cotransfection of p175 alone gave a very low level of activation (track 3). However, cotransfection with both p175 and p111 resulted in a high level of gD promoter activity as measured by detection of correctly initiated RNA (track 4). Activation by cotransfection with p111 alone was not detectable in this assay (track 12). These results are consistent with earlier observations (Everett, 1984). The efficiencies of activation by cotransfection of p175 with insertion or deletion mutants of Vmw110 were very similar to those expected on the basis of the CAT assay results. Insertions in region 1 (p110F1, p110E8-1 and p110E32-1; tracks 5–7) were all low in activity compared with p111, while the deletion in region 3 (p110E18; track 8) was even less active. An insertion in region 5 (p110E2; track 11) was also greatly reduced. The insertion in region 4

Fig. 7. Quantitation of the amount of RNA derived from the gD promoter after activation by cotransfection with normal and mutant IE genes. **Lanes 1 and 13** contain mol. wt markers (pBR322 cut by *HpaII*) with the 147 and 160 nucleotide bands indicated. **P** marks the position of the full length probe. **Lanes 2–12** contain S1-nuclease-resistant DNA after hybridization of total cytoplasmic RNA from transfected cells to a probe which detected RNAs derived from the gD promoter, as described in the text. HeLa cells were transfected with pRED4 alone (**track 2**) or cotransfected with p175 (**track 3**) or p111 (**track 12**). **Lanes 4–11** show the results of analysis of RNAs derived from cells transfected with pRED4 and p175 (in all cases) plus p111, p110F1, p110E8-1, p110E32-1, p110E18, p110E17-1, p110E1-1 or p110E2 respectively. RNA correctly initiated at the gD promoter gives a characteristic pattern of bands around the 160-nucleotide marker (Everett, 1983).

(p110E17-1; track 9) was, as expected, not as severely reduced as the others. An insertion which was equally active as p111 in the CAT assays (p110E1-1) gave reduced levels of RNA in this experiment (track 10) but this result is within the limits of variation observed for a single experiment. Therefore, the conclusion from these data is that the CAT assay results given in Table I are consistent with the phenotypes of the mutations as measured by direct quantitation of the RNA (Figure 7).

Discussion

This paper describes a detailed functional analysis of Vmw110, a *trans*-acting transcriptional regulator of HSV-1, which illustrates

that at least five regions of the polypeptide contribute to its function. Alteration of the DNA sequence of a gene within its coding region could, in theory, affect gene function at a variety of different levels. The stability, processing and transport of the RNA could be altered, as could its secondary structure, which might impede translation. In that none of the mutations are within sequences known to affect these aspects of RNA metabolism, the interpretation of the results assumes that these factors are not involved in the explanation of the observed effects. Indeed, the results of the immunofluorescence experiments show that a protein related to Vmw110 is produced from all the mutated genes described. Change in the primary sequence of a polypeptide might affect its function in a variety of ways. The most direct would be the disruption of an 'active' site or a binding site for some essential co-factor. Alternatively, the alteration might affect processing, such as phosphorylation, of the polypeptide so as to reduce its activity. Changes in secondary and tertiary structure of the protein induced by insertions or deletions far from active or modification sites might disrupt protein shape so as to indirectly affect function at these sites. The cellular localization or transport of the polypeptide could be affected, and lastly the mutant protein may be less stable, leading to a lower intracellular concentration. None of the mutations appeared to affect cellular localization, and there were no extreme variations in the amount of nuclear fluorescence seen in each case, but it is not possible to state which of the other possible explanations applies to any of the mutants described here. However, it is clear that a number of specific regions of the polypeptide are important for its function and that the experimental method used is a promising approach to investigate the mechanism of action of Vmw110 and other *trans*-acting transcriptional activators.

It is also possible that changes in the amino acid sequence could lead to an increase in the activity of the protein. This would probably be a relatively rare event, and there is little indication that 'improved' versions of the polypeptide have been created. Those insertions giving a higher activity than the parent show only small increases that are probably not significant.

While the analysis of the results has concentrated on the position of the disruption in the primary sequence of the protein, it is probable that the actual nature of the inserted amino acids could also play a role in the phenotype of the mutant. The sequences of the insertions are given in Table I. Many have a net basic charge which could obviously alter the properties of the polypeptide. However, there appears to be a much greater correlation between the position of an insertion and the phenotype rather than the exact nature of that insertion; the insertions resulting in a reduction in activity fall into groups (as discussed above), while insertions of a given sequence can have a small or large effect depending on their position. Analysis of the effect of the nature, rather than the position of the alteration of the primary sequence, would be best achieved by site-directed mutagenesis of selected residues. One of the values of this work is that it provides a rational basis on which to select amino acids for specific alteration.

The assay used in this study utilizes clones of IE genes 1 and 3 from HSV-1 strain 17⁺, a CAT expression vector linked to the HSV-1 gD promoter and HeLa cells. Under these conditions the activation of gene expression by either Vmw110 or Vmw175 by themselves is slight, while in combination the two are highly synergistic. Other studies using IE genes cloned from different strains of HSV, different indicator promoters and different cell types have recorded somewhat different findings including higher activities of Vmw175 and Vmw110 by themselves and a less pronounced synergism (O'Hare and Hayward, 1985a,b; Quinlan and

Knipe, 1985; Gelman and Silverstein, 1985). It is an intriguing question whether the phenotypes of the mutants observed in this study might differ in some aspects using another assay system. In particular, is the phenotypic pattern of the mutants the same in gene activation assays in the absence of Vmw175? This question is currently under study.

Computer analysis of the primary sequence of Vmw110 has revealed a number of interesting features (Perry *et al.*, 1986). The region between amino acids 130 and 160 shows homology both to an open reading frame in the VZV DNA sequence that lies in a comparable region of the VZV genome (3' to the Vmw175 homologue) and also to a cysteine-rich motif that has been implicated in binding metal ions and DNA (Berg, 1986). Deletion mutant p110de11 has lost the whole of this region while p110F1 contains an insertion within it. Both mutations produce a polypeptide that is located in the nucleus with a normal fluorescence pattern (data not shown) and both (particularly p110F1) retain some activity in the presence of Vmw175 (Tables I and II). Therefore this region is not absolutely essential for activity in this assay (although p110de11 is severely reduced) but it is interesting that insertions and deletions in the 150 bp 3' to the cysteine-rich region exhibit significant loss of activity (Table I; region 1, Figure 5a). The section of the protein from around amino acids 240–550 was noted as being especially G–C rich and therefore contains a predominance of the amino acids alanine, glycine, proline and arginine (Perry *et al.*, 1986). It has been suggested that such highly G–C-rich regions do not reflect a requirement for the encoded amino acids for function but that the extreme base composition results from other (unknown) mechanisms which produce sequences which are tolerated by the protein (McGeoch *et al.*, 1986). In that only two of the 19 insertions in this region show activities of < 50%, it appears that in general the highly G–C-rich region is not critical for protein function. However, p110E28-1 maps near the centre of this area and has a very low level of activity. In addition, deletion mutants within this region (region 3, Figure 5a; p110E18 and p11050F) are also poor activators. Therefore, in this instance [and also perhaps in the region (region 2) of the insertion in plasmid p110E31-1] a region of the protein encoded by highly G–C-rich DNA is important for function, but it may be that mutations around amino acid 400 affect the general shape of the polypeptide rather than an important 'active site'. The region (region 4) from amino acids 555–591 is extremely rich in serine residues which might be possible sites of phosphorylation. The three insertion mutations in this area all show moderate loss of activity; it would be interesting to examine the extent of phosphorylation of these and other mutants.

Hydropathicity and secondary structure plots of the Vmw110 sequence have been computed (L.J. Perry and D. McGeoch, personal communication). These predict that the protein is rather hydrophilic with extensive areas of 'random coil' and 'extended side chain' interspersed by short regions of alpha helix and 'reverse turn' structures. However, there are no striking correlations between these features and the positions of any of the mutants with deleterious effects on activity. Of the two regions which contain the most severely affected mutations, region 1 between amino acids 180 and 220 is predicted to contain some alpha helix and two relatively hydrophilic regions, while region 5 between amino acids 630 and 720 is mostly in the 'extended side chain' structure and of average hydrophilicity.

The mode of action of Vmw110 in the infected cell is not known. It may increase transcription by interacting directly with the DNA or by interaction with other viral polypeptides (possibly

Vmw175) or with cellular transcription factors. A study of the regions of the polypeptide that are important for function is one method of approaching this problem. It is now possible to transfer the most interesting of the mutants (such as p110E2 and p110E8-1) into the viral genome to study the virology of such lesions and to use the mutant viruses as vectors with which to purify the mutant polypeptides. Comparison of the physical properties of the mutant and wild-type proteins (such as DNA binding and post-translational modification) will lead to important insights into the mechanism of action of this transcriptional activator.

Materials and methods

Plasmids

Plasmid pgDCAT has been described before (Everett, 1986). Plasmid p111 was derived from p110de17 (Perry *et al.*, 1986) by 'filling-in' the *EcoRI* site in the vector multi-cloning region. Plasmid p175 contains an HSV-1 fragment extending from the *BamHI* site at position +27 in the IE gene 3 transcribed region, through the 'a' sequence, to the *SstI* site 900 bp into IR_L. The SV40 early promoter and enhancer region from the *HpaII* site to the *HindIII* site (SV40 coordinates 5171–346) has been joined to the IE gene 3 coding region using a *BamHI* linker. Plasmid pRED4 has been described before (Everett, 1983).

Construction of insertion mutants

Plasmid p111 was cut with restriction enzymes *HaeIII*, *RsaI*, *AvaI* or *FspI* in the presence of ethidium bromide so as to produce a maximum of singly cut linear molecules. The precise conditions depended on the enzyme used, but generally 10 µg of plasmid was cut with 10 units of enzyme for 30 min at 37°C in the presence of 20–50 µg/ml ethidium bromide in a final volume of 50 µl of the buffer recommended for the enzyme. The DNA was loaded onto a 5–20% sucrose density gradient and centrifuged for 16 h at 30 000 r.p.m. in the Sorvall TST41 rotor. Linear DNA sedimented more slowly than open circular DNA such that the two were easily separated. DNA that had been cut with enzymes giving a blunt-ended cut was ligated directly to 12-bp *EcoRI* linker oligonucleotides. *AvaI*-cut DNA was first 'filled-in' with Klenow DNA polymerase before ligation to 8-bp *HindIII* linker oligonucleotides. Plasmids containing linker insertions were screened by analysis of small-scale plasmid preparations. The locations of insertion in *RsaI*, *AvaI* and *FspI* sites was ascertained by detailed restriction mapping, while those in *HaeIII* sites were mapped first using restriction enzymes and the locations were confirmed by DNA sequence. The number of linker oligonucleotides inserted was deduced from restriction mapping; in those cases where multiple linkers had been inserted the plasmids were extensively digested with the enzyme cutting the linker sequence, re-ligated and again tested by restriction enzyme analysis. Deletion mutants were constructed or isolated as described in the text and their extents determined by DNA sequence analysis.

Transfection of HeLa cells and CAT assays

HeLa cells were seeded at 1.0×10^6 cells/50 mm Petri dish and transfected with 4 µg of each plasmid DNA by the calcium phosphate precipitation method (Corsalo and Pearson, 1981). pUC9 was used to equalize the amount of DNA in comparative experiments when different numbers of plasmids were used. After 24 h the cells were washed and fresh medium added. After a further 24 h, extracts were prepared and CAT activities assayed as described by Gorman *et al.* (1982). The radioactivity in the substrate and monoacetylated product (position 3) spots were determined by scintillation counting to calculate the percentage conversion of substrate to product. The protein concentration of each extract was determined (Lowry *et al.*, 1951) and the percentage conversion per mg protein calculated. This value was then expressed as a percentage of that obtained in a control transfection containing pgDCAT, p175 and p111. Generally, one-thirtieth of the extract obtained from a 50-mm dish was incubated with 4 nmol [¹⁴C]chloramphenicol for 30 min at 37°C. In a typical experiment, ~25% acetylation of the substrate was obtained in the positive control transfection, giving specific activities of the order of 300 pmol acetylated/µg extract/h.

Immunofluorescence

HeLa cells were seeded into 50 mm Petri dishes and transfected with 10 µg of p111 or mutant plasmids as described above. After 24 h the cells were washed and re-fed. A further 24 h later, the cells were washed twice with phosphate-buffered saline (PBS) containing 1% calf serum and then fixed with 2% formaldehyde in PBS with 2% sucrose for 10 min at room temperature. After washing three times with PBS/1% calf serum, the cells were permeabilized with 0.5% NP40 in PBS/1% calf serum/10% sucrose for 5 min at room temperature, and again washed three times with PBS/1% calf serum. The plates were then dried for a few minutes at 37°C, and selected areas marked for staining. The cells

were then incubated for 1 h at room temperature, firstly with a rabbit antiserum directed against an oligopeptide from the carboxy terminal end of the Vmw110 protein (Perry *et al.*, 1986), and then with fluorescein-conjugated sheep anti-rabbit Ig antiserum. After each incubation the cells were washed three times with PBS/1% calf serum and then dried. Finally the cells were examined by u.v. fluorescent microscopy.

Preparation and analysis of total cytoplasmic RNA

The methods used have been described in detail before (Everett, 1983). Total cytoplasmic RNA was prepared by extensive phenol/chloroform extraction of NP40 lysed cells and ethanol precipitation. RNA (40 µg) was hybridized to a 5'-end labelled single-stranded probe isolated from pRED4. The hybrids were treated with S1 nuclease, analysed on an 8% acrylamide gel containing 7 M urea and detected by autoradiography.

Acknowledgements

The author thanks Mr S. McCallum and Mrs M. Wilkie for excellent technical assistance, Dr M. Frame for the rabbit anti-Vmw110 oligopeptide antiserum, Dr D. McGeoch for interesting discussions and Professor J. H. Subak-Sharpe and Dr N. Stow for helpful criticism of the manuscript.

References

- Ackermann, M., Braun, D.K., Pereira, L. and Roizman, B. (1984) *J. Virol.*, **52**, 108–118.
- Barany, F. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4202–4206.
- Berg, J.M. (1986) *Science*, **232**, 485–487.
- Clements, J.B., Watson, R.J. and Wilkie, N.M. (1977) *Cell*, **12**, 275–285.
- Corsalo, C.M. and Pearson, M.L. (1981) *Somatic Cell Genet.*, **7**, 603–616.
- Everett, R.D. (1983) *Nucleic Acids Res.*, **11**, 6647–6666.
- Everett, R.D. (1984) *EMBO J.*, **3**, 3135–3141.
- Everett, R.D. (1985) *EMBO J.*, **4**, 1973–1980.
- Everett, R.D. (1986) *J. Gen. Virol.*, **67**, 2507–2513.
- Gelman, I.H. and Silverstein, S. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5265–5269.
- Gorman, C.M., Moffatt, L.F. and Howard, B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
- Harland, J. and Brown, S.M. (1985) *J. Gen. Virol.*, **66**, 1305–1321.
- Hay, R.T. and Hay, J. (1980) *Virology*, **104**, 230–234.
- Honess, R.W. and Roizman, B. (1974) *J. Virol.*, **14**, 8–19.
- Jones, P.C. and Roizman, B. (1979) *J. Virol.*, **31**, 299–314.
- Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) *J. Biol. Chem.*, **193**, 265–275.
- McGeoch, D.J., Dolan, A., Donald, S. and Brauer, D.H.K. (1986) *Nucleic Acids Res.*, **14**, 1727–1764.
- O'Hare, P. and Hayward, G.S. (1985a) *J. Virol.*, **53**, 751–760.
- O'Hare, P. and Hayward, G.S. (1985b) *J. Virol.*, **56**, 723–733.
- Pereira, L., Wolff, M.H., Fenwick, M. and Roizman, B. (1977) *Virology*, **77**, 733–749.
- Perry, L.J., Rixon, F.J., Everett, R.D., Frame, M.C. and McGeoch, D.J. (1986) *J. Gen. Virol.*, **67**, 2365–2380.
- Poffenberger, K.L., Tabares, E. and Roizman, B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2690–2694.
- Preston, C.M. (1979) *J. Virol.*, **29**, 275–284.
- Quinlan, M.P. and Knipe, D.M. (1985) *Mol. Cell. Biol.*, **5**, 957–963.
- Sacks, W.R. and Schaffer, P.A. (1987) *J. Virol.*, **61**, 829–839.
- Sacks, W.R., Greene, C.C., Aschman, D.P. and Schaffer, P.A. (1985) *J. Virol.*, **55**, 796–805.
- Sears, A.E., Halliburton, I.W., Meignier, B., Silver, S. and Roizman, B. (1985) *J. Virol.*, **55**, 338–346.
- Stow, N.D. and Stow, E.C. (1986) *J. Gen. Virol.*, **67**, 2571–2585.
- Wagner, E.K. (1985) In Roizman, B. (ed.), *The Herpesviruses* Vol. 3, Plenum Press, pp. 45–104.
- Watson, R.J. and Clements, J.B. (1980) *Nature*, **285**, 329–330.

Received on November 4, 1986; revised on April 10, 1987