Identification of a novel multifunctional structural domain in the herpes simplex virus type 1 genome: implications for virus latency

J. P. Quinn,1 R. A. McGregor,1 C. E. Fiskerstrand,1 C. Davey,2 J. Allan2 and R. G. Dalziel1

Departments of Veterinary Pathology1 and Biochemistry2, University of Edinburgh, Edinburgh EH9 1QH, UK

A domain, previously termed RE1, exists within the herpes simplex virus type 1 genome potentially influencing expression of immediate early genes and the latency associated transcripts. This domain consists of 10 tandem copies of a CT-rich sequence. We demonstrate that this domain binds multiple host-cell factors that may allow RE1 to act either as a transcriptional regulator and/or to affect nucleosomal and DNA structure in the latent genome.

Herpes simplex virus type 1 (HSV-1) can establish a latent infection that persists for the lifetime of the host. Latent virus genome exists in an extrachromosomal state and is complexed with host nucleosome protein (Mellerick & Fraser, 1987; Deshmane & Fraser, 1989). This chromosomal structure may have consequences for the establishment, maintenance of, and reactivation from latency. The area encompassing the latency associated transcripts (LATs) and the immediate early (IE) genes is implicated in reactivation and we have analysed this region of the HSV-1 genome to determine whether it contains domains which might regulate transcription over this large area. CpG motifs in eukaryotic genes are often found clustered 5’ of the transcription start site and can regulate both the transcriptional and methylated state of the loci (Cross & Bird, 1995). Our analysis revealed extensive CpG motifs with one notable exception being a domain lying 5’ of the LAT promoter. This domain represents the tandem repeat region previously termed reiteration set 1 (RE1) (Perry & McGeoch, 1988) (Fig. 1A). The sequence of the element constituting this repeat is (CCCCCTCTCCCCCTCT) × 10. We predicted that this unit was capable of binding to factors that can alter chromatin structure or act as transcription regulators. Factors potentially recognizing and binding to this motif are shown in Fig. 1(B). In all cases the effects could act over a large distance and affect several genes near this locus. Although proteins expressed during viral infection could further regulate this region, we chose initially to determine whether neuronal host-cell proteins could interact in a sequence-specific manner with this element. Host proteins are likely to be involved, at least in part, in regulating maintenance or reactivation from latency.

Binding of factors to this domain was analysed by electrophoretic mobility shift analysis (EMSA). Nuclear extract was prepared (Quinn et al., 1989) from either the brains of adult Wistar rats or HeLa cells which were grown in Dulbecco’s modified Eagle’s media (DMEM) supplemented with 10% heat-inactivated calf serum. Each reaction using double-stranded oligonucleotide contained 0.5 ng 32P 3’-end-labelled oligonucleotide probe generated using Klenow enzyme. This end-labelled probe was mixed with 20 µg protein extract which had been pre-incubated on ice with 1 µg poly(dI–dC):(dI–dC). Competitor oligonucleotide where appropriate was added to the extract at the same time as the probe. Each reaction using single-stranded oligonucleotide contained 0.5 ng 32P 5’-end-labelled oligonucleotide probe generated using T4 kinase. All probes were purified over G25 spin columns. Reactions were incubated for 15 min at room temperature and separated on a 4:2% polyacrylamide (29:1) gel. As RE1 is a multimerized element we chose to address binding to one of the repeats. EMSA demonstrated that the motif would interact with three classes of transcription factor.

1 Classical double-stranded DNA-binding proteins recognizing CT-rich elements (Fig. 2A)
2 Single-stranded DNA-binding proteins (Fig. 2B, C)
3 Nucleosomal phasing proteins related to BGP1 (Fig. 3 A, B)

Growth factors, including nerve growth factor (NGF), are known to modulate HSV-1 latency in model systems (Wilcox & Johnson, 1988; Wilcox et al., 1990). dG-rich motifs are growth factor-regulated in the gastrin gene (Bachwick et al., 1992) and in general dGC-rich elements can mediate the action of growth factors, e.g. NGFIA (Kendall et al., 1994). The HSV-1 GC motifs within RE1 may therefore act as transcriptional regulators under appropriate conditions during virus infection and latency in sensory neurons. Enhancer elements in viruses are often found to be multimerized motifs which support higher levels of gene expression than the single element alone.
often acting synergistically rather than additively (Herr & Clarke, 1986; Quinn et al., 1989).

An enhancer function in a virus would not be unexpected; however, the two additional properties (2 and 3 above) of this domain may represent a novel mechanism involved in alphaherpesvirus latency. These additional properties would also change the topology over this region having subsequent consequences for the parameters regulating latency. Firstly, EMSA demonstrated that each of the single strands of a single RE1 repeat bound in a sequence-specific manner to nuclear proteins. Single-stranded DNA-binding proteins have been demonstrated to act as transcriptional regulators (DavisSmyth...
genes to allow transcription has been observed most elegantly in the mouse mammary tumour virus model (Archer et al., 1992). To investigate BGP-1 interactions with RE1, BGP1 was affinity purified from adult chicken erythrocyte nuclear extracts as previously described (Clark et al., 1990; Lewis et al., 1988) and used in EMSA (Fig. 3). Purified BGP1 did not bind directly to RE1; however, complementing this protein with HeLa cell nuclear extract (at a level that does not itself generate any specific complex) demonstrated sequence-specific complex formation on the RE1 elements. We have previously observed such a precedent for complementation of transcription factor-binding activity using a purified fraction and extract (Quinn et al., 1989). The binding of factors exemplified by hnRNP K and BGP1 to the DNA could have distinct and disparate consequences for chromatin/nucleosomal structure of the virus genome. Indeed, binding of double-stranded sequence-specific factors or BGP-related proteins might be expected to significantly antagonize the action of the single-stranded DNA-binding proteins and vice versa.

A recent publication has demonstrated that single-stranded regions are found in vivo to be associated with genes rescheduled for reactivation after mitosis (Michelotti et al., 1997). Obviously, these single-stranded regions are therefore acting as genetic markers to control expression patterns. Although latency occurs within post-mitotic neurons such a genomic marker in the HSV-1 genome might allow for initiation of appropriate temporal gene expression. Alternatively, such a region may account for the expression of the LAT transcript observed within a sub-population of neurons containing latent HSV-1, by acting to keep the genome accessible to transcription factors. The potential complexity of this RE1 domain could be increased by the action of nucleosome restructuring by such as BGP1 and virus-encoded proteins.

Our data indicate that the RE1 domain may have an important and novel regulatory function. The location of this domain indicates that it could affect expression of both LAT and IE and transcripts. Alternatively, as regulatory elements can affect transcription over a very large genomic fragment, this domain could affect expression in other HSV-1 genomic locations. Consistent with this, single-stranded regions generated by the action of single-stranded DNA-binding proteins might facilitate the interaction of factors over areas that otherwise might be energetically unfavourable, by acting in a manner similar to a hinge within the promoter (Michelotti et al., 1996, 1997). The number of copies of the repeat element, each capable of interacting with transcription factors, suggests that it could have a strong regulatory effect. Recently, the DR2 repetitive element within the HSV-1 genome has been shown to have regulatory properties (Martin & Weber, 1998). Interestingly, these authors suggested that DR2 might have distinct properties in a genomic structure compared to that present within a plasmid used in transient transfection analysis. It will be of interest to analyse cell-specific interactions with
this motif in dorsal root ganglia and the variation in binding in response to virus infection or stimuli that result in reactivation of HSV-1 latent genomes. Ultimately, the characterization of the function of this domain will require construction of recombination viruses deleted in both sets of the RE1 repeat and analysis in vivo. The potential for such domains as a general regulatory mechanism in alphaherpesvirus latency is indicated both by the HSV-1 DR2 data and sequence analysis which reveals similar motifs in the related varicella-zoster virus (VZV), in the reiterated element R4. The R4 domain is located between two IE genes that are transcribed in latently infected cells (Davison & Scott, 1986), perhaps suggesting a functionally similar genomic function in both viruses.

An MRC short-term project grant to J.P.Q. and R.G.D. and a BBSRC award to J.P.Q. supported this work.

References


Received 27 April 1998; Accepted 2 July 1998