Sequence analysis of the herpes simplex virus type 1 strain 17 variants 1704, 1705 and 1706 with respect to their origin and effect on the latency-associated transcript sequence

Fazil Junejo, Alasdair R. MacLean and S. Moira Brown*

MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, U.K.

The precise endpoints of the deletions/insertions in three variants (1704, 1705 and 1706) of herpes simplex virus type 1 (HSV-1) strain 17 have been determined by dideoxynucleotide sequence analysis. The analysis was undertaken to discover whether the three variants had arisen from the same initial event and the extent of the deletions with respect to the latency-associated transcripts (LATs) and the proposed LAT promoter region. It is not possible from the deletion boundaries to determine unequivocally whether the three variants had arisen from the same recombination event although 1706 could be descended from 1705 by illegitimate recombination. The results demonstrate that spontaneous deletions can occur at random within RL, the extent of the deletions in Ur is constrained by the essential nature of Ur genes in vitro but is otherwise arbitrary and deletions in 1704 completely remove both copies of the LAT promoter region and in IRr extend into the 5' end of the LAT sequence.

Herpes simplex virus (HSV) establishes latency in neurons of sensory ganglia from which it reactivates intermittently throughout life to produce lesions at peripheral sites. Only recently has some understanding of the molecular basis of latency become apparent with the identification of latency-associated transcripts (LATs) (Stevens et al., 1987; Spivack & Fraser, 1987). During latency, transcription is limited to the long repeat region (Rl) of the genome and three transcripts of 2.0, 1.5 and 1.45 kb can be detected both by Northern blot hybridization and in situ hybridization. These transcripts are detected both in ganglia from animal model systems and in ganglia explanted from human cadavers. In addition, transcripts from regions, adjacent to the three LATs have been detected by in situ hybridization although not by Northern blotting. These transcripts have been designated the minor LATs (mLATs) (Mitchell et al., 1990).

We previously reported the isolation and characterization of three deletion variants (1704, 1705 and 1706) of HSV type 1 (HSV-1) strain 17 (MacLean & Brown, 1987). The variant 1704 has a 3.8 kb deletion in IRL and the adjacent portion of UL plus a 1.2 kb deletion in TRL. The variant was able to become established, was maintained and was reactivated from latency using the mouse eye model system. However, following ganglion explant it was found that the reactivation of 1704 was significantly delayed compared to its parental 17+ strain. LAT transcripts could not be detected either by in situ hybridization or Northern blotting in the ganglia of mice infected with 1704. It was therefore concluded that LATs play a role in the timing of reactivation from explant culture (Steiner et al., 1989). Further analysis showed that sequence elements necessary for initiating transcription or stabilizing mLATs are deleted in the variant 1704 (Mitchell et al., 1990). Based on the position of the LAT promoter region proposed by Weschler et al. (1988), it was determined by Southern blotting and restriction enzyme analysis that the TRl deletion in 1704 removed the promoter but did not extend into the LAT transcript, whereas the IRr deletion removed the promoter and the 5' end of the LATs.

This paper describes dideoxynucleotide sequence analysis of the three variants (1704, 1705 and 1706) isolated from the same transfection to determine the precise position of the deletions with respect to the LAT promoter and transcripts, and whether the three isolates could have arisen from the same recombination event. The HpaI v*r*r* fragment (Fig. 1, lane 2) spanning the U1/IRr deletion in 1704 (MacLean & Brown, 1987) was cut from a 1% agarose gel, electroeluted and purified through a Sepharose column (Maniatis et al., 1982). The fragment was cloned into the Smal site of pUC19 and subcloned into the BamHI and EcoRI sites of M13mp18 and M13mp19 (Sanger et al., 1980). Sequencing was carried out using the chain termination reaction method of Sanger et al. (1977), and Klenow polymerase I and [32P]dATP were used for preparing radiolabelled DNA.
Short communication

(a) HpaI

(b) BamH1

Fig. 1. (a) HpaI and BamH1 restriction maps of HSV-1 strain 17. (b) Autoradiographs of HpaI (lanes 1 to 3) and BamH1 (lanes 4 to 5) digests of viral DNA labelled with 32P in vivo. Lanes 1 and 4, HSV-1 strain 17; lane 2, 1704; lane 3, 1705; lane 5, 1706. Novel fragments are indicated (*).

Table 1. Hybridization of oligonucleotides to the HpaI o* fragment of 1704

<table>
<thead>
<tr>
<th>Oligonucleotide coordinates (nucleotide position)</th>
<th>Hybridization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 6291 to 6307</td>
<td>+</td>
</tr>
<tr>
<td>2. 7041 to 7057</td>
<td>+</td>
</tr>
<tr>
<td>3. 8054 to 8070</td>
<td>-</td>
</tr>
<tr>
<td>4. 8721 to 8737</td>
<td>+</td>
</tr>
<tr>
<td>5. 9491 to 9507</td>
<td>+</td>
</tr>
</tbody>
</table>

DNA products were resolved on 6% acrylamide gels containing 9 M-urea, and after electrophoresis the gels were dried and subjected to autoradiography.

Dideoxynucleotide sequencing determined the precise extent and coordinates of the deletion. The deletion is 3758 bp in length and spans the UL/IRL junction starting at nucleotide 116502 and ending at nucleotide 120260. The deletion removes 655 bp of UL and 3103 bp of IR, starting 1232 bp downstream of the 3' end of IE1. Only 170 bp of the 3' end of UL56 is retained. The region 799 bp from the 5' end of the LATs is deleted as well as the LAT promoter region (TATAA, CAAT and SpI binding sites) at 681 to 959 bp upstream of the 5' end (Fig. 3).

An initial estimate of the position of the variant 1704 TR deletion was obtained by Southern blot hybridization (Southern, 1975) using five synthetic oligonucleotides ranging from nucleotides 6291 to 9507. Only one failed to hybridize (Table 1) indicating that the deletion was between nucleotides 7057 and 8720. To confirm this, sequence analysis was carried out on the novel HpaI o* fragment (Fig. 1, lane 2) previously shown to span the deletion (MacLean & Brown, 1987). Sequencing was initiated using oligonucleotide number 2 (Table 1) as a primer and it was found that the deletion was 942 bp in length extending from nucleotides 7202 to 8144, which are entirely within TR. The LAT transcript is not affected but the promoter region from nucleotides 7592 to 7870 is removed (Fig. 3).

The novel s*/r* HpaI fragment spanning the deletion in variant 1705, (Fig. 1, lane 3) was cloned and sequenced as for 1704. The deletion was found to be 4735 bp in length extending from nucleotides 115453 to 120188. The deletion removes 3031 bp of IR and 1704 bp of UL, completely excising UL55 and UL56, and stops 183 bp and 694 bp downstream of the 3' ends of IE2 and IE1, respectively. Variant 1705 is not deleted in TR (Fig. 3) although 714 bp of the 5' end of the LATs and the promoter region have been lost.
In the variant 1706, sequences deleted from the right-hand end of UL have been replaced by sequences from the left-hand end (MacLean & Brown, 1987). A novel BamHI fragment b*/c* (Fig. 1, lane 5) comprising sequences from the b fragment and the c fragment was cloned and digested with BamHI/HpaI/SmaI to determine the subfragment spanning the deletion. Southern blots in which BamHI b and c were hybridized to digested DNA identified the composite DNA in a 726 bp subfragment (Fig. 2). Sequencing of this fragment revealed a 1807 bp deletion at the right-hand end of UL which had been replaced by 4755 bp from the left-hand end. The deletion started 80 bp downstream of the 3' end of IE2 at nucleotide 115350 and terminated at nucleotide 117157. As in the variant 1705, the deletion completely removed UL55 and UL56. The sequence from nucleotides 9215 to 13970 has been repeated and inserted in an inverted orientation to replace the deletion. Consequently 1706 contains two copies of UL1, UL2, UL3 and UL4, and a second partial copy of UL5 (Fig. 3).

A diagram illustrating in detail the structures of 1704, 1705 and 1706 around the deletions and insertion is shown in Fig. 3.

The three variants 1704, 1705 and 1706 with UL/IRL/TRL rearrangements were isolated following restriction enzyme analysis of 80 progeny plaques from a single recombination infection. Among the same progeny, 11 isolates showed extensive variation up to several hundred base pairs long within the Rl regions of the genome excluding the ‘a’ sequence (MacLean & Brown, 1987). It was apparent therefore that these progeny molecules contained a higher than expected proportion of genomes with alterations and rearrangements involving the long repeat region of the genome. This was in agreement with findings on the HSV-2 strain HG52, from which several variant isolates with rearrangements of the short region of the genome were isolated following a single infection, whereas variants with long region alterations arose from a separate experiment (Brown & Harland, 1987; Harland & Brown, 1985). This clustering of rearrangements to particular regions of the genome in any one experiment suggests they arose from a single or related events. The explanation at the time of isolation of 1704, 1705 and 1706 was that one variant genome had arisen which, due to instability, went through several rounds of rearrangement prior to the formation of stable genomes; thus one initial event could potentially lead to the isolation of several variants in a population.

To test this hypothesis the altered regions of the genomes in 1704, 1705 and 1706 were subjected to dideoxynucleotide sequence analysis. The variant 1704 is deleted between nucleotides 116502 in UL and 120260 in IRl. Therefore 6112 bp of IRl has been retained between the ‘a’ sequence and the start of the deletion; 3103 bp has been lost in addition to the 655 bp of UL. In TRl, 7202 bp has been retained between the ‘a’ sequence and the start of the deletion in addition to 1071 bp between the deletion and the start of UL. The TRl deletions had therefore not arisen from the same point and are not within or immediately adjacent to any of the four sets of tandem reiterations located in the long repeat region of HSV-1 (Perry & McGeoch, 1988).

The 4734 bp deletion in 1705 started at nucleotide 115453 in UL and terminated at nucleotide 120118 in IRl. There was therefore 1049 bp between the start of the UL deletions in 1704 and 1705. Within IRl there was only a 142 bp difference between the endpoints of the 1704 and 1705 deletions. Again the IRl deletion in 1705 is not connected with the reiterated subsets. 1704 and 1705 may therefore have arisen from a common process involving a nicked DNA strand and recombination event traversing to different points in the genome.

In 1706 there is no RL deletion and the deletion in UL extends to nucleotide 115453 which is 103 bp to the left of the 1705 deletion point. A model for the origin of 1706 involving recombination between 1705 in the prototype orientation and either 174 or 1705 in the Ir orientation was proposed by MacLean & Brown (1987). This depended on the UL/IRl novel junction being unstable and therefore prone to disruption. The closeness of the two endpoints in 1705 and 1706 would not refute this model. Sequence comparisons have demonstrated no gross homologies between the sequences around nucleotides 115350 and 13970, which is the insert point of the DNA from the left-hand side of UL.
Fig. 3. Diagram showing the positions of the deletions in the HSV-1 strain 17 variants 1704 and 1705, and the deletion/insertion in 1706. Co-ordinates are given in nucleotide positions. Unique and adjacent RL sequences are illustrated, and the co-ordinates of the UL/RL and U5/R5 junctions are given. The positions of the 5' ends of the LATs, the LAT promoter regions and the 3' ends of IE1 in TR5 and IR5 are indicated. The extents of the deletions are shown (D) and the positions of the 1706 insert and relevant genes are illustrated.

The model proposed for the expansion or contraction of the repeats (McGeoch, 1984; Whitton & Clements, 1984) depends on the essential nature of the genes in UL adjacent to RL. UL55 and UL56 are non-essential and UL54 is essential (Sacks et al., 1985). The termination point of a deletion could therefore be entirely arbitrary within the sequences between the 3' end of UL54 and the UL/IR5 junction, or it may depend on the structure and conformation of the DNA with particular regions being more accessible. Our sequence analyses of 1704, 1705 and 1706 deletions cannot distinguish between these two possibilities.

It is evident that 1704, 1705 and 1706 (i) do not have deletions with precisely the same endpoints, (ii) the IR5 deletions of 1704 and 1705 could be related, (iii) RL deletions are not dependent on tandem reiterations, (iv) 1706 could have arisen by illegitimate recombination as proposed, (v) UL deletions (repeat extension) are controlled by the essential nature of the genes (genomes with UL1 deletions never having been isolated) and (vi) the deletion start/stop point within non-essential DNA is probably arbitrary.

When Steiner et al. (1989) demonstrated that 1704 failed to produce LATs which could be detected by
Northern blotting or in situ hybridization, in addition to reactivating slowly from latency, the precise extent of the deletions with respect to the LATs and their promoter region had not been determined; it was assumed that the absence of detectable LATs could only be due to the deletions affecting the transcripts and/or promoter region. Sequence analysis of 1704 has shown that in Ul/IRL, 170 bp of UL56 has been retained and the deletion does not affect IE1, the 3' end of which is at nucleotide 120882. The 5' ends of the LATs are at nucleotide 119461 which means that 794bp of the deletion does not affect IE1, the 3' end of which is at nucleotide 119461 which means that 794bp of the transcripts has been deleted. Weschler et al. (1988) showed that the LAT promoter region was located between 681 and 959bp upstream of its 5' end (nucleotides 118502 to 118780) and is therefore totally absent from the IR<sub>L</sub> region of 1704. In TR<sub>L</sub> the LAT transcript, the 3' end of which is at nucleotide 6911, is not removed but the deletion from nucleotides 7202 to 8144 completely removes the promoter region at nucleotides 7592 to 7870. 1704 therefore has no LAT promoters but retains one complete copy of the LATs and two-thirds of the other copy.

In conclusion, this analysis demonstrates that spontaneous deletions probably occur at random within the long repeats and are not dependent on reiterated sequences. The isolation of a genome in which 4 kb has been deleted from R<sub>L</sub>, including both complete copies of IE1 (unpublished results) substantiates this conclusion. The extent of deletion into Ul is constrained by the essential nature of the Ul genes in vitro but otherwise appears to be arbitrary.

We are grateful to John Subak-Sharp for his interest and support and to June Harland for her willing assistance. F.J. is a Pakistan Government Scholar.

References


(Received 7 March 1991; Accepted 21 May 1991)