Nucleotide sequence analysis of a homologue of herpes simplex virus type 1 gene US9 found in the genome of simian herpes B virus

A. M. Killeen,* L. Harrington, L. V. M. Wall† and D. C. Kelly

Chemical and Biological Defence Establishment, Porton Down, Salisbury, Wiltshire SP4 0JQ, U.K.

The 10K gene of simian herpes B virus (SHBV) has been located and the nucleotide sequence determined. Its relationship to homologous genes in other herpesviruses has been examined. The SHBV 10K gene exhibits a closer relationship to its homologue in HSV-1 than to those in the other viruses studied. Nucleotide sequence identity of 61% was found between the HSV-1 and SHBV 10K genes, and 57% identity was found between the corresponding predicted protein sequences. A comparison of the amino acid sequences of the herpesvirus 10K proteins has revealed a number of conserved features. These are examined in relation to possible functions of the 10K proteins. Implications for evolutionary relationships between SHBV and other herpesviruses are discussed.

Simian herpes B virus (SHBV) [or B virus, as it was called after the initials W.B. of the first patient (Sabin & Wright, 1934)] is an alphaherpesvirus which causes a mild, recurring disease in its natural hosts, monkeys of the macaque genus. However, symptomatic SHBV infection in man, by accidental monkey-bite or from infected tissues or fluids, is a neurotropic illness which is fatal in almost 80% of cases. The most recent reported example of B virus disease in humans occurred in Michigan, U.S.A. in 1989 (Katz et al., 1990). Three people became infected with SHBV with one subsequent fatality.

In monkeys, SHBV causes an infection similar to herpes simplex infection in man. Primary infection is characterized by both oral and genital lesions. These symptoms clear spontaneously in 7 to 14 days but primary infection is followed by a life-long latent infection from which the virus may reactivate to produce a recurrence of the disease (Hartley, 1966). Studies of SHBV began in the 1960s when serological tests showed a close relationship between SHBV and herpes simplex virus type 1 (HSV-1) (Plummer, 1964; Norrild et al., 1978). The striking difference in pathogenesis of these closely related viruses for the human host challenges investigators to discover the evolutionary relationships of SHBV and HSV. Examination of the molecular biology of SHBV may help solve this hitherto unanswered question. During a study of SHBV glycoproteins, a small open reading frame (ORF) was found in the region of a SHBV genome which had been sequenced. In this communication we identify this ORF as another herpesvirus homologue of HSV-1 US9. The sequence data presented in this paper represent the first evidence of the arrangement of genes in SHBV.

SHBV DNA used in this work was from the strain of virus designated prototypic B virus (Wall et al., 1989). The prototypic strain was an oral isolate from a cynomolgus monkey (Vizozo, 1975). SHBV was grown in Vero cells and total DNA (cellular and viral) was isolated from SDS-phenol-extracted cultures according to the method of Lonsdale (1979). The resultant crude DNA extract was used in this work. Similarity to HSV-1 US9 was localized in the SHBV genome using a DNA fragment containing the HSV-1 US9 gene as a probe. A plasmid library of cloned KpnI fragments of HSV-1 (Glasgow strain 17) in pAT153 (Davison & Wilkie, 1983), was a kind gift from J. Oram, Centre for Applied Microbiological Research, Porton Down, Salisbury, Wiltshire, U.K. The plasmid clone containing the HSV-1 fragment designated KpnI h was digested with BamHI to release a 1.8 kbp fragment (designated BamHI z; Davison & Wilkie, 1983), containing the coding region for the HSV-1 US9 gene. BamHI z was labelled with [3²P]dCTP (Amersham) by extension of random hexamer primers (Feinberg & Vogelstein, 1985) and used to probe the SHBV genome using the technique of Southern blotting. The HSV-1 US9 gene probe was allowed to hybridize to filters overnight at 65 °C. Filters were washed at high stringency, allowing 30% mismatch
between probe and target sequence, before autoradiography overnight at -70 °C. Binding to the probe was obtained from a plasmid library of phage M13. Restriction mapping of the SHBV genome by sequencing both strands of DNA. The entire restriction digests of SHBV DNA which had been subjected to electrophoresis through 0-8% agarose gels, the aforementioned plasmid library were isolated from sequences. Due to its ability to bind to HSV-1 US9 in DNA cross-hybridization studies and in its nucleotide sequence, we conclude that ORF2 encodes a homologue to other herpesvirus genes. ORF2 is a homologue of HSV-1 US9. ORF3 is in the opposite orientation to ORF1 and ORF2, is incomplete and encodes a homologue of HSV-1 US10.
Fig. 2. Relationships between five herpesvirus 10K protein homologues. (a) Multiple amino acid identity alignment between homologous proteins from SHBV (SHBV10K), HSV-1 (HSV1US9), PRV (PRV11K), EHV-1 (EHV1-ORF2) and VZV (VZVUS1).

The EHV-1 sequence used here corresponds to the terminal 110 amino acids of ORF2 which encodes a total of 220 amino acids (Cullinane et al., 1988). Sequences were aligned using the CLUSTAL suite of programs (Higgins & Sharpe, 1988). Asterisks (*) indicate residues exactly conserved across all five sequences. Dots (.) indicate residues where conservative substitutions occur. (b) Dendrogram showing possible phylogenetic relationships among the herpesvirus 10K protein homologues. The multiple amino acid alignment from (a) was used to produce the dendrogram using the CLUSTAL programs.

of HSV-1 US9. Hereafter, this will be referred to as the SHBV 10K gene. In HSV-1, the US9 gene encodes a phosphoprotein which is found in the tegument of virions. It is also found associated with nucleocapsids in the nuclei of infected cells (Frame et al., 1986). As the HSV-1 protein has an Mr of 10026, it has been designated 10K (McGeoch et al., 1985).

The 90 amino acid sequence encoded by the SHBV 10K gene (Fig. 1) produces a protein with a predicted Mr of 9.93K. There are ten strongly basic amino acids (lysine and arginine) and 10 strongly acidic amino acids (aspartic acid and glutamic acid). The overall charge on the protein at pH 7.0 is predicted to be -0.016.

Homologues of HSV-1 US9 have been found in the 11K gene of pseudorabies virus (PRV) (Petrovskis et al., 1987) and the US1 gene of varicella-zoster virus (VZV) (McGeoch et al., 1985). The carboxy-terminal half of the 23K protein encoded by equine herpesvirus type 1 (EHV-1) ORF2 (Cullinane et al., 1988) also shows significant resemblance to the HSV-1 10K protein. All four genes encoding 10K protein homologues are located in the short unique component of their respective genomes. A multiple alignment of the SHBV 10K protein and its homologues in HSV-1, PRV, VZV and EHV-1 is presented in Fig. 2(a). A number of features are evident from this alignment. Firstly, although EHV-1 ORF2 encodes a protein of 23K, the carboxy-terminal half of the protein shows significant identity to the 10K proteins. The function of EHV-1 ORF2 has not been investigated but it is possible that the extra domain at the amino terminus may have a distinct function.

A striking feature is the number of potential sites for phosphorylation in the proteins. This is not surprising as the HSV-1 homologue is a phosphoprotein in which serine and threonine constitute 16% of its amino acids (Frame et al., 1986). Serine and threonine residues form 13% of the amino acids of the SHBV 10K protein, and 14% of the amino acids of the PRV 11K and VZV US1 proteins. The portion of the EHV-1 23K protein which shows homology to the 10K proteins (i.e. the carboxy-terminal 110 amino acids) has a serine plus threonine content of 15%. All five proteins exhibit conservation of two tyrosine residues (Fig. 2a, positions 59 and 60). Although tyrosine-linked phosphorylation is not as commonly found in viral proteins as serine- or threonine-linked phosphorylation, the complete conservation of the two tyrosine residues indicates that tyrosine-linked phosphorylation may occur at these sites.

A cluster of arginine residues is found in the SHBV 10K amino acid sequence at positions 83 to 87 (Fig. 2a). A similar sequence is also found in the HSV-1 10K, PRV 11K and EHV-1 23K proteins. The basic arginines are replaced in the VZV 10K protein by two lysines and a single arginine at this position. It has been suggested that such a hydrophilic section represents a basic nuclear accumulation sequence (Frame et al., 1986; for a review of such signals see Dingwall & Lasky, 1986). This hypothesis is supported by the finding that the HSV-1 10K protein is localized in the nucleus of infected cells.

The function of the HSV-1 10K protein and its homologues is at present unknown. HSV-1 mutants which have the 10K gene deleted from the genome remain viable in cell lines (Umene, 1986). PRV 11K gene deletion mutants retain infectivity but lose virulence (Lomniczi et al., 1984). Similarity to a number of protein kinases (PKs) and other ATP-binding enzymes found in database searches at the nucleotide and protein levels suggests that the HSV-1 10K protein may have a role in
phosphorylation. The HSV-1 10K protein and its homologues all contain many potential sites for phosphorylation. Some PKs are known to autophosphorylate. A number of different PK activities appear to be involved in the phosphorylation of viral and cellular proteins following herpesvirus infection (Stevely et al., 1985). Owing to the difficulties in distinguishing virus-induced PK activities from those encoded by the host cell, the number, origin and substrate specificities of PKs associated with herpesvirus-infected cells is not known. In order to study this further, the kinase activity of the purified 10K protein should be investigated in a cell-free system.

As the HSV-1 10K protein has been found in the nucleus of infected cells (Frame et al., 1986), and clusters of basic amino acids have been found in its homologues, the potential DNA-binding activity of the 10K proteins should be investigated. Frame et al. (1986) found that the HSV-1 10K protein does not bind to calf thymus DNA-cellulose indicating that it is unlikely to be a DNA-binding protein. However, closer examination of the positions of the acidic and basic residues of the SHBV 10K protein reveal that 100% of the negatively charged residues (aspartic acid and glutamic acid) occur in the amino-terminal half of the protein, whereas 90% of the positively charged residues (lysine and arginine) occur in the carboxy-terminal half of the protein. The same pattern is seen in the amino acid sequences of the HSV-1 10K, PRV 11K and VZV 11K proteins, and in the EHV-1 homologue. There is therefore potential for the arginine-containing carboxy-terminal portion to bind DNA in a manner analogous to DNA binding protein II of Escherichia coli (Rouviere-Yaniv & Yaniv, 1979). This E. coli protein is composed of two 9.5K chains which associate in pairs as dimers. The molecule has extended arginine-containing arms which interact with the phosphates of one turn of the DNA backbone. It has been suggested that dimerization of the HSV-1 10K protein could account for some of the higher M, forms which have been observed (Frame et al., 1986). Computer searches of nucleotide and protein databases for homology to the SHBV 10K gene have revealed homologous sequences in a number of DNA-binding proteins. Speculation about the possible interaction of the 10K proteins with DNA in the nucleus of infected cells can only be resolved by repeated DNA-binding assays. Analysis of the structure of the protein from crystallographic data or computer-generated predictions of molecular structure could also help to elucidate this matter.

The nucleotide sequence of 670 bp of SHBV DNA upstream of the 10K gene was determined and studied (Fig. 1). An ORF of 66 amino acids was found to extend upstream of the sequenced DNA (ORF1). ORF1 was found to show nucleotide sequence identity to HSV-1 glycoprotein E (gE) at a level of 56-2%. Alignment of the amino acid sequence encoded in the incomplete ORF1 with that of the HSV-1 gE protein was not possible owing to insufficient amino acid similarity between the sequences. From the evidence presented here, it cannot be concluded that ORF1 encodes a homologue of the HSV-1 gE gene. Further work is necessary to identify this gene conclusively.

Examination of the nucleotide sequence downstream of the SHBV 10K gene revealed the presence of a large partial ORF encoded on the opposite strand to the 10K gene (Fig. 1, ORF3). Computer alignment of the SHBV DNA sequence with the sequence of HSV-1 US showed identity of 63.4% between ORF3 and HSV-1 US10, which encodes a polypeptide of 33K (McGeoch et al., 1985). HSV-1 US10 contains 936 bp. As expected, the incomplete ORF3 (622 bp) aligns along the carboxy-terminal two-thirds of HSV-1 US10. An identity of 50.3% was found between the HSV-1 33K protein sequence and the predicted protein encoded by ORF3. From these results it may be concluded that ORF3 is a homologue of HSV-1 US10; it will hereafter be referred to as the SHBV 33K gene.

The 33K protein encoded by HSV-1 US10 is a basic virion protein of unknown function (McGeoch et al., 1985). US9 and US10 are adjacent and encoded in the opposite orientation to each other in the HSV-1 genome. The sequencing data presented here demonstrate identical organization for the SHBV homologues of US9 and US10.

An evolutionary tree based on the amino acid sequences of the 10K homologues of SHBV, HSV-1, PRV, EHV-1 and VZV is presented in Fig. 2(b). The resultant evolutionary pattern supports the hypothesis that SHBV is the simian counterpart of HSV and therefore more closely related to HSV than to the other herpesviruses (Sabin, 1934). In addition, the evidence presented in this study supports previous findings for evolutionary trees of the herpesvirus which depict divergent pathways for HSV, for PRV and EHV-1, and for VZV. This pattern of evolution has been predicted from computer analyses of 12 different herpesvirus thymidine kinase sequences (Griffin & Boursnell, 1990) and seven different herpesvirus glycoprotein B sequences (Whalley et al., 1989). It is clear from the study presented here that SHBV should now join HSV on its branch of the evolutionary tree.

We wish to thank Keith Martin and Bryan Lingard for assistance with computer analyses.

References


