The nucleotide sequence of the equine herpesvirus 4 thymidine kinase gene

Lesley Nicolson, Ann A. Cullinane and David E. Onions

1Department of Veterinary Pathology, University of Glasgow Veterinary School, Bearsden Road, Bearsden, Glasgow G61 1QH, U.K. and 2Irish Equine Centre, Johnstown, Naas, County Kildare, Eire

We have determined the genomic location and nucleotide sequence of the equine herpesvirus 4 thymidine kinase (TK) gene. The gene is positioned at approximately 0.48 map units within the long unique component of the genome and is flanked by genes encoding a herpes simplex virus 1 (HSV-1) UL24 homologue and glycoprotein H. The predicted protein is composed of 352 amino acids, has an $M_r$ of 38,800 and exhibits 36% identity to the predicted TK of HSV-1.

Equine herpesviruses types 1 and 4 (EHV-1 and -4) are alphaherpesviruses responsible for significant economic losses within the equine industry. EHV-1 induces abortion, respiratory disease and neurological disease. EHV-4, previously classified as EHV-1 subtype 2, is primarily associated with respiratory disease although EHV-4-induced abortions are reported occasionally (Allen & Bryans, 1986).

During lytic infection, many herpesviruses induce a virus-specific deoxypyrimidine kinase (thymidine kinase, TK), an enzyme operative in the salvage pathway of pyrimidine biosynthesis (Kit, 1985a). TK proteins of distinct viral origins differ in their affinity for natural substrates such as thymidine and deoxycytidine and for nucleoside analogues such as acyclovir, which, in a phosphorylated form, inhibit the DNA replicative cycle of specific herpesviruses (Kit, 1985a; Kit et al., 1987). TK− herpesviruses are often less neurovirulent than their TK+ counterparts due to a reduced capacity to replicate within resting cells in vivo (Field & Wildy, 1978; Kit, 1985a). TK gene function has also been linked to herpesvirus pathogenicity (Kit et al., 1985a, b) and to the capacity to reactivate from a latent state (Efstathiou et al., 1989; Tenser et al., 1989).

The nucleotide sequences of the TK genes of a variety of herpesviruses, including EHV-1, have been reported (reviewed in Robertson & Whalley, 1988; Honess et al., 1989; Mittal & Field, 1989; Scott et al., 1989). Here we report the nucleotide sequence of the EHV-4 TK gene and compare selected regions of the predicted amino acid sequence of its protein to those of other herpesvirus TK proteins.

The EHV-4 genome is a dsDNA molecule of 144 kbp, composed of a long (L) and short component (S), the latter bounded by a set of inverted repeats (Cullinane et al., 1988). The EHV-4 TK gene was expected to map at around 0.5 map units (m.u.) given the reported position of the EHV-1 TK gene (Robertson & Whalley, 1988) and evidence for collinearity of the long components of the EHV-1 and EHV-4 genomes (Cullinane et al., 1988). Southern blot analyses were performed in which recombinant plasmids containing EHV-4 BamHI fragments mapping between 0.39 and 0.62 m.u. were digested with restriction endonucleases and hybridized to a 5.35 kbp BamHI/ClaI probe derived from the left end of EHV-1 BamHI fragment B (Fig. 1). This probe contains the entire TK gene of EHV-1 and flanking sequence (Robertson & Whalley, 1988). Restriction digest products were electrophoresed through 0.75% agarose and transferred to a GeneScreen membrane (New England Nuclear). The membrane was baked for 2 h at 80°C, prehybridized overnight at 42°C in 45% formamide, 4.2 x SSC, 4 x Denhardt's solution, 8% dextran sulphate, 0.04 M-PO$_4^-$, 80 µg/ml salmon sperm DNA, 0.08% SDS and hybridized for a similar period following the addition of 2 x 10$^7$ c.p.m. of denatured, $^{32}$P-labelled EHV-1 TK DNA. The blot was washed in 2 x SSC, 0.1% SDS at 65°C, dried and exposed overnight. The EHV-1 probe hybridized to the BamHI C fragment of EHV-4 (0.43 to 0.53 m.u.), to a 10 kbp BamHI/ClaI fragment (BC10) and to three SmalI subfragments of C mapping between 0.48 and 0.53 m.u. (Fig. 1, 2a, 2b). The ability of plasmids containing BamHI C and subfragment BC10 to transform BHK TK− cells to a TK+ phenotype confirmed the presence of a functional TK gene within these fragments. To localize the TK gene further, subfragments of EHV-4 C were cloned into a Bluescript M13+ plasmid vector and 20 µg of each construct was transfect-
ed into monolayer BHK TK− cells by a modification of the technique of Graham & van der Eb (1973). TK+ colonies were selected in HAT-supplemented medium (hypoxanthine 10⁻⁴ M, aminopterin 4 × 10⁻⁵ M, thymidine 1.6 × 10⁻⁵ M). TK-transforming activity was thus localized to a 2 kbp EcoRV/XhoI fragment (RX2), cloned in construct pBSRX2, with a map position of approximately 0.48 (Fig. 2b).

The nucleotide sequence of both strands of fragment RX2 and adjoining DNA was determined by using single-stranded plasmid DNA as template and Bluescript-derived and custom-made oligonucleotides as primers in a Sanger dideoxynucleotide sequencing strategy (Sanger et al., 1977) (Fig. 1c). Analysis was performed using Beckman Microgenie software (Queen & Korn, 1984). Three open reading frames (ORFs) were detected, one in the lower strand (ORF1) and two in the upper strand (ORF2 and ORF3). ORF1 encodes a 272 amino acid polypeptide with 36% identity to the predicted products of gene UL24 of herpes simplex virus type 1 (HSV-1) (McGeoch et al., 1988) and gene 35 of varicella-zoster virus (VZV) (Davison & Scott, 1986). The product of the HSV-1 UL24 gene is important for viral growth in vitro (Jacobson et al., 1989). ORF2 is the only intact ORF within RX2, the smallest tested fragment with TK-transforming activity. It encodes a 352 amino acid product which shares 36% identity with the HSV-1 TK gene product. ORF3 extends for 2565 bp and encodes the EHV-4 glycoprotein H (gH) gene (Nicolson et al., 1990). The gene order UL24 homologue-TK-gH, conserved throughout most herpesvirus genomes, is thus conserved in EHV-4.

TATA box homologues (Corden et al., 1980), TTATA and TATTTAA, are located 85 and 305 bp upstream of the initial methionine codon of ORF1 (at 865 bp and 645 bp in Fig. 3). Deletion of 525 bp of DNA proximal to the EcoRV site of RX2 (181 to 705 in Fig. 3) by exonuclease III digestion of pBSRX2 did not adversely affect TK-transforming activity of the recombinant plasmid (authors' unpublished observations). We suggest, therefore, that the TATA box positioned at bp 861 to 865 (Fig. 3) is the functional TATA box of the EHV-4 TK gene. A potential RNA polymerase initiation site (Corden et al.,...
The predicted EHV-4 TK initiation codon is positioned at 949 to 951 bp within the sequence GTAATGG. The coding region of the gene commences at 949 bp and extends to a TGA termination codon at 2005 bp. A poly(A) signal AATAAA (Proudfoot & Brownlee, 1976), is positioned 42 bp downstream of the termination codon.
The translated product of the EHV-4 TK gene is a 352 amino acid protein of predicted Mr, 38,800. Comparison of its predicted amino acid sequence to those of other herpesvirus TKs (Wagner et al., 1981; Kit et al., 1983; Baer et al., 1984; Otsuka & Kit, 1984; Kit, 1985a, b; Davison & Scott, 1986; Robertson & Whalley, 1988; Honess et al., 1989; Mittal & Field, 1989; Scott et al., 1989; Sheppard & May, 1989) indicated identities as follows: 28 and 29% with marmoset herpesvirus (MarHV) TK and bovine herpesvirus 2 (BHV-2) TK, 35 to 37% with the TKs of HSV-1, HSV-2, VZV and BHV-1, 42% with pseudorabies virus (PRV) TK, and 89% with EHV-1 TK. Comparison to gammaherpesvirus TKs indicated identities of 24% with Epstein–Barr virus (EBV) TK (amino acids 208 to 527) and herpesvirus saimiri (HVS) TK (amino acids 283 to 607) and of 33% and 34% with Marek’s disease virus (MDV) TK and herpesvirus of turkeys (HVT) TK respectively; the latter two seem to be more similar to alphaherpesvirus than to gammaherpesviruses in terms of their molecular biology.

Multiple alignments of herpesvirus TKs (Robertson & Whalley, 1988; Honess et al., 1989; Mittal & Field, 1989; Scott et al., 1989) have highlighted several regions of conserved sequence, two of which are detailed in Fig. 4. The most N-terminal of these conserved regions (Fig. 4a) shares identity with a sequence domain within many nucleotide-binding proteins and has been proposed to form part of a nucleotide-binding site (Gentry, 1985; Otsuka & Kit, 1984). The predicted EHV-4 TK nucleotide-binding domain conforms to the consensus sequence GXXGXXGKT derived from functional and comparative studies (Otsuka & Kit, 1984; Gentry, 1985; Liu & Summers, 1988), differing, as for EHV-1, PRV and MarHV, only in the substitution of a serine residue for threonine. A second sequence region conserved between herpesvirus TKs corresponds to the putative nucleoside-binding site (Fig. 4b). Darby et al. (1986) have proposed that the catalytic centre of the HSV-1 TK enzyme, active as a dimer, is composed of three distinct regions of polypeptide arranged to form overlapping nucleoside- and nucleotide-binding sites. In addition to amino acids within the conserved nucleotide- and nucleoside-binding domains, a cysteine residue at position 336 forms part of the catalytic centre. Like most other herpesvirus TKs, the EHV-4 TK possesses a cysteine residue [Cys (315)] at a comparable position. The precise role of other conserved regions of the herpesvirus TKs with respect to the structural and functional integrity of the protein remains to be determined.

The determination of the precise genomic position of the EHV-4 TK gene is a prerequisite to the production, by recombinant DNA techniques, of EHV-4 mutants with defined TK deletions. Studies on the growth characteristics and latency capacity of such mutants should lead to a greater understanding of the role of TK in herpesvirus pathogenesis.

We would like to thank J. M. Whalley and G. R. Robertson for making available to us unpublished sequence data. Thanks also to M. Riddell, A. May and J. Fuller for their help in the preparation of the manuscript. We are grateful to the Horserace Betting Levy Board and the Equine Virology Research Foundation for their generous financial support.

References

Short communication


(Received 29 March 1990; Accepted 24 April 1990)