

GENE 03438

Characterization of the gene and an antigenic determinant of equine herpesvirus type-1 glycoprotein 14 with homology to gB-equivalent glycoproteins of other herpesviruses

(Recombinant DNA; unusual signal sequence; hydropathic analysis; membrane-anchoring region; epitope)

Peixuan Guo

Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, NY 12201 (U.S.A.) and Laboratory of Virus Diseases, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, MD 20892 (U.S.A.)

Received by J.A. Engler: 6 March 1989

Revised: 21 August 1989

Accepted: 23 August 1989

SUMMARY

The gene encoding glycoprotein 14 (gp14) of equine herpesvirus type 1 was sequenced. Nucleotide sequence analysis revealed a complete transcription unit composed of a CAT box, a TATA box, a ribosome-binding sequence, a polyadenylation signal and an open reading frame (ORF) of 2940 bp transcribed from left to right. The amino acid (aa) sequence deduced from this ORF corresponded to that of a protein with 979 aa and had the characteristic features of membrane gp including a 20-aa signal sequence at the N terminus, a 743-aa surface domain, a 40-aa membrane anchoring region, a 108-aa hydrophilic cytoplasmic domain at the C terminus and eleven potential sites for N-linked glycosylation. An unusual feature of this protein was an exceptionally long (66 aa) sequence, with a preponderance of hydrophilic residues, preceding the hydrophobic signal core. An antigenic determinant recognized by an anti-gp14 monoclonal antibody was present in the N terminus of the postulated surface domain. Comparison of gp14 with the gp of other herpesviruses indicated that gp14 was highly homologous to corresponding gp of pseudorabies (gII), bovine herpesvirus (gI), varicella-zoster virus (gII), as well as of herpes simplex virus, Epstein-Barr virus and human cytomegalovirus (gB).

INTRODUCTION

Equine herpesvirus type 1 (EHV-1), a member of the α -herpesvirus subfamily, is an infectious agent of horses causing acute upper respiratory tract infection, contagious viral abortion and neurological disease. The biological and

biochemical properties of the EHV-1 closely parallel those of human herpes simplex virus. Thus, EHV-1 exhibits the typical characteristics of herpesviruses, including genomic isomerization, establishment of latent infections, induction of neurological disorder, generation of defective interfering particles, in vitro oncogenic transformation of cells, and α - β - γ gene regulation (O'Callaghan et al., 1981). EHV-1 has been utilized as a model system for studying the biological consequences of herpesvirus infection.

Glycoproteins (gp) are the principal immunogens of herpesvirus, stimulating both humoral and cellular immune response in the host. The well-characterized gp of HSV include gB, gC, gD, gE, gG, gH and gI. It has been demonstrated that gB and gD can elicit neutralizing antibodies from the infected host (Paoletti et al., 1984; Cantin et al., 1987). gC could stimulate class-I restricted CTL (Rosenthal et al., 1987; Glorioso et al., 1985) and gD could

Correspondence to: Dr. P. Guo, Department of Veterinary Pathobiology, Purdue University, West Lafayette, IN 47907 (U.S.A.)
Tel. (317)494-7543; Fax (317)494-0781.

Abbreviations: aa, amino acid(s); β Gal, β -galactosidase; BHV-1, bovine herpesvirus type 1; bp, base pair(s); CTL, cytotoxic T lymphocytes; EBV, Epstein-Barr virus; EHV-1, equine herpesvirus type 1; gB, glycoprotein B; gp, glycoprotein(s); gp, gene encoding gp; HCMV, human cytomegalovirus; HSV-1, herpes simplex virus type 1; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); ORF, open reading frame; PRV, pseudorabies virus; Sig, signal sequence(s); VZV, varicella-zoster virus.

TACACCGTTGAACCGTGGTGTACCGACTCTCAAGAGACTAGTCGTTTATGATAACTGCGGC	62	AGCCCATAGTCTCCGCCAACCCCTTGACGAGCGAGTTCGACGAGGGTCTCTGGGGGACGTTG	2222
TAAGGTGAATGGTCAATAGCGAAGTTTCAAGGTTTATTTGCTTGAAGGGAGTGAC	122	SerAlaIleValSerAlaThrLeuAspGluArgValAlaAlaArgValLeuGlyAspVal	641
AGGTGTGACGGCCACGCGCGCGCTGGCTGGTAAATATACGGGGAGCTACTTAGCCGCG	192		
GCAGTATTCCTCGGTTTCCACTGGGAGGATGAGTGGCTCTCTCGCGCCACATCGTACCTA	242	ATAGCTATAACGCACTGCGCCAAAATAGAGGGCAACGTGTACTTGCAAACTCCATCGCGC	2282
CCCGGACTCCGCGCCACAGTGTGCTGGTGGAGTGCATTACATAACCTACGAGGCGTCACT	302	IleAlaIleThrHisCysAlaLysIleGluGlyAsnValThrLeuGlnAsnSerMetArg	661
Met	1		
TCCTCTGGTTCGCGTCTCTGTCGCGGCTCCACATGGGCAATTTGGCGGGAGACGGTGT	362	TCGATGGACAGTAACAGTGTACTCCGCCCCCCCCAACATTACAACTTAAGAAT	2342
SerSerGlyCysArgSerValGlyGlySerThrTrpGlyAsnTrpArgGlyAspGlyGly	21	SerMetAspSerAsnThrCysTyrSerArgProProValThrPheThrIleThrLysAsn	681
GATTACGACAGGACGTGTCTCTCTCTGTATGTCAGTGTCTGCTCAGCAGCTGGCTCTGG	422	GCAANCAACAGAGGGTCGATAGAAGCCAGCTGGGAGGAGAGACAGATTTTTCACGGAG	2402
AspLeuArgGlnArgArgValLeuSerProValCysSerAlaProAlaAlaGlySerTrp	41	AlaAsnAsnArgGlySerIleGluGlyGlnLeuGlyGluGluAsnGluIlePheThrGlu	701
ATCGGGAGCAACTAGGCAATGTGGAACTTACTCGCCACCCCGCCGCTGGGAAAG	482	CGCAAGCTGATCGAGCGCTGCGCCCTCAATCAGAAGCGCTACTTTAAGTTTGGCAAGAG	2462
IleGlySerGlnLeuGlyAsnValGlyAsnLeuLeuAlaThrProHisProLeuGlyLys	61	ArgLysLeuIleGluProCysAlaLeuAsnGlnLysArgTyrPheLysPheGlyLysGlu	721
CCGGCATCAGAGGTTGGGCACAATAGTTTACGCTGTTTGTGCTTTTGGAAAGCTGT	542	TACGTTTACTACGAGAATACACGTTCTGTCGCAAGTGCCTCCACGGAAATCGAGTT	2522
ProAlaSerSerArgValGlyThrIleValLeuAlaCysLeuLeuLeuPheGlySerCys	81	TyrValTyrTyrGluAsnTyrThrPheValArgLysValProProThrGluIleGluVal	741
GTGTTAGAGCGTACCACCCAGCCCAAGCCCACTAGTACTCCCACTTCCATGTCA	602	ATCAGCAGTACGTTGAACTAACTTACGCTTTTGGAAAGCCGCGAGTTTCTGCCCTG	2582
ValValArgAlaValProThrThrProSerProProThrSerArgThrProThrLysMetSer	101	IleSerThrTyrValGluLeuAsnLeuThrLeuLeuGluAspArgGluPheLeuProLeu	761
ACGCATCCCATGGGACAGTACAGCTACGCTCCCAACAGAAACCGCCGACCCACTC	662	GAGGTGTACAGCGGGCTGAGCTGGAGACACCGCCCTGCTAGACTACAGCGAAATACAG	2642
ProHisSerHisGlyThrValAspProThrLeuLeuProThrGluThrProAspProLeu	121	ArgProTyrThrArgAlaGluLeuGluAspThrGlyAlaValSerLysThrValSerGlyIle	781
AGACTGGCTGTGCGGAGTCCGATATACTCGTGGAGTGGAGACTTTTACACTGCCCA	722	CGCCGACACAGCTCCAGCTCTCAGTTTTTACGACATCGACGCGTGTCAACGTGGAC	2702
ArgLeuAlaValArgGluSerGlyIleLeuAlaGluAspGlyAspPheTyrThrCysPro	141	ArgArgAsnGlnLeuHisAlaLeuArgPheTyrAspIleAspSerValValAsnValAsp	801
CCGCTCCCGGATCCACCGTCTGACGCAACACCTAGAACTTGCCTCAAGTTTGGAC	782	AATCCCGAGTATTATCAGGGATCGCCAGCTTTTCAAGGGCTGGGTAAGTGGGGG	2762
ProHisThrGlySerThrValValArgIleGluProProArgThrCysProLysPheAsp	161	AsnThrAlaValIleIleArgGlySerProAlaPheSerArgAlaTrpValLysTrpGly	821
CTGGGGAACCTTACCGAGGGGATGCTGTTATTTTAAAGAAACATCGCTCCCTAC	842	AGCCGCTGGGAACGCTGTTCTCGCGCGCGCGCTGTTGTTTCAACCGTATCTGGAATA	2822
LeuGlyArgAsnPheThrGluGlyIleAlaValIlePheLysGluAsnIleAlaProTyr	181	ArgProTrpGluArgSerPheSerAlaArgGlyAlaValSerLysThrValSerGlyIle	841
AAATTCAGGGCAACGTATACACAAGACATCGTTGTAACAGTGTGGAAAGGATAC	902	GCTTGTCTTTTAAACAACCCATTTGGGGGCTAGCCTCGCCTGCTGGTAAATCGCCGC	2882
LysPheArgAlaAsnValTyrTyrLysAspIleValValThrArgValTrpLysGlyTyr	201	AlaCysPheLeuAsnAsnProPheGlyGlyLeuAlaIleGlyLeuValIleAlaGly	861
AGCCATACGCTCCCTCCGACAGATACAAATGACAGGGTTCGGTTCGGTGGAGGAGAT	962	CTGGTAGCTGCGTTTTTTGCTTACAGATATGTAATGCAGATCCGAGTAAACCCATGAAA	2942
SerHisThrSerLeuSerAspArgTyrAsnArgValProValSerValGluGluIle	221	LeuValAlaAlaPhePheAlaTyrArgTyrValMetGlnIleArgSerAsnProMetLys	881
TTGCGTCTACGACAGTAAAGGAAATGTTCTGCAAGGCCGAGTACCTCAGAGATAAC	1022	GCTCTATACCCCATAAACAAGGCTTGAAMAAACAAGCCAACTTCTCAGCGCCAG	3002
PheGlyLeuIleAspSerLysGlyLysCysSerSerLysAlaGluTyrLeuArgAspAsn	241	AlaLeuTyrProIleThrThrLysAlaLysLysAlaLysLysLysAlaIleLysLys	901
ATCATGCCACCGCTACCACGACGAGGACGAGGTGGAGCTTGTATTTGGTCCGCTCC	1082	AACGAGGAGGACGATGGGAGCGACTTTGATGAGGCCAAGCTGAAGGGCTCGCAATG	3062
IleMetHisHisAlaTyrHisAspAspGluAspGluValGluLeuAspLeuValProSer	261	AsnGluGluAspAspGlySerAspPheAspGluAlaLysLeuGluGluAlaArgGluMet	921
AAAGTTGCACTCCGGGGCCAGAGCTGGCAGACCCAAACGATACTACGCTTACGTT	1142	ATCAAATACATGCTATGGTTTCGGCCCTGGAAAAGCAGGAAAAGAAAGCTATAAAGAAA	3122
LysPheAlaThrProGlyAlaArgAlaTrpGlnThrThrAsnAspThrThrSerTyrVal	281	IleLysTyrMetSerMetValSerAlaLeuGluLysGlnGluLysLysAlaIleLysLys	941
GGTGGATGCCATGGAGGCACTACACGCTCAAGCTGTGCAACTGATCGAGGAGGTT	1202	AACAGTGGGTTGGCTGATCCCGCACTAACGCTCAAAGCTGGCCCTGCGAAGCCGCT	3182
GlyTrpMetProTrpArgHisTyrThrSerThrSerValAsnCysIleValGluGluVal	301	AsnSerGlyValGlyLeuIleAlaSerAsnValSerLysLeuAlaLeuArgArgArgGly	961
GAGCGCGGTCGCTACCCCTACGACTCTCCGCTGTCACCGGTGATATGTGTAC	1262	CCCAAATATACCGACTCCAACGAGACGATCCATGGAAAATGAAATAATGGTTAAACA	3242
GluAlaArgSerValTyrProTyrAspSerPheAlaLeuSerThrGlyAspIleValTyr	321	ProLysTyrThrArgLeuGlnGlnAsnAspThrMetGluAsnGluLysMetVal	979
CGCTTCCGTTTTACGGCTGAGGCTCGCCCTCGCATAGACCAATAGCTACCGCAG	1322	TGTTTAAATAATATATGACACGACTCAAAGTGTGACCTCATATTTGCATAACCACCTC	3302
AlaSerProPheTyrGlyLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	341	TAGTTCGGGCCAAAGGATATTTAAGCTAGTATCTCCGCCGAAAGG	3345
GAGCGTTTACGCAAGTTAAGGGTACAGCCCGCGACTTAGACAGTAACTACAAGCC	1382		
GluArgPheArgGlnValGluGlyTyrArgProArgAspLeuAspSerLysLeuGlnAla	361		
GAAAGCGGTTTACCAAAAATTTTATCACTACCCCGCATGTCACCTCAGCTGGAATCG	1442		
GluGluProValThrLysAsnPheIleThrThrProHisValThrValSerTrpAsnTrp	381		
ACCGAGAAGAAAGTCCGAGCGGTGACGTCGACCAATGGAAGAGGTCGACGAACTCGCT	1502		
ThrGluLysLysValGluAlaCysThrLeuThrLysGluValValAspGluLeuVal	401		
AGGACGAGTTCGCGGGTCTCAGACTTACTATTCGATCCATCTCGTCTACTTTATC	1562		
ArgAspGluPheArgGlySerTyrArgPheThrIleArgSerIleSerSerTyrPheIle	421		
AGTAACTACTCAATTAAGTTGAAAGTGCCTTACTGAAATGTGTATCCAAGAA	1622		
SerAsnThrThrGlnPheLysLeuGluSerAlaProLeuThrGluCysValSerLysGlu	441		
GCAAAGGAAGCCATAGACTCGATATACAAAAGCAGTACGAGTCTACGCACTCTTAGC	1682		
AlaLysGluAlaIleAspSerIleTyrLysLysGlnTyrGluSerThrHisValPheSer	461		
GGTATGTGGAATATTACCTGGCAGCGGGGGTCTTAATGCAATTCAGACCTATGCTC	1742		
GlyAspValGluTyrTyrLeuAlaArgGlyGlyPheLeuIleAlaPheArgProMetLeu	481		
TCCAACTCGCCAGGCTGACTGACAGGCTGTGTGAGATCTAACCGCACCTACGAC	1802		
SerAsnGluLeuAlaArgLeuTyrLeuAsnGluLeuValArgSerAsnArgThrTyrAsp	501		
CTAAAAAATCTATTGAACCCCAATGCAACAATAACAATCAACCCAGCGAAGACGCAAG	1862		
ThrAlaLysLeuLeuAsnProAsnAlaAsnAsnAsnAsnThrThrArgArgArgArg	521		
TCTCTCTGTCAGTACCAGAACCTCAGCCAAACCAAGATGGTGTGATAGAGAACAAAT	1922		
IleLeuLeuSerValProGluProGlnProThrGlnAspGlyValHisArgGluGlnIle	541		
CTACATCGCTTGCACAACAGCAGTGGAGCAACCGCAGGTACCGATTCTTCCAACGTC	1982		
LeuHisArgLeuHisLysArgAlaValGluAlaThrAlaGlyThrAspSerSerAsnVal	561		
ACCGCCAAACAGCTGGAGCTCAAAAACCCAGCTGCTATGAGTTTCCATGCTACAG	2042		
ThrAlaLysGlnLeuGluLeuIleLysThrThrSerSerIleGluPheAlaMetLeuLeu	581		
TTGCATACGATCACATCCAATCCACGCTAATGAAATGCTAAGTAGAATAGCACTGCG	2102		
PheAlaTyrAspHisIleGlnSerHisValAsnGluMetLeuSerArgIleAlaThrAla	601		
TGGTGTCCCTCAAACAAGAGCGGCCCTATGGAACGAAATGGTGAAGATTACCCCG	2162		
TrpCysProLeuGlnAsnLysGluArgProLeuTrpAsnGluMetValLysIleThrPro	621		

Fig. 1. The nt sequence of the gene encoding gp14 of EHV-1 and the deduced aa sequence. Arrows indicate the CAT box and the TATA box of the promoter, as well as the polyadenylation signal, AATAAA. The signal sequence at the N terminus and the membrane-spanning region at the C terminus are underlined. The predicted N-X-T/S consensus domains for N-linked glycosylation are marked with asterisks. Bold-face letters indicate the epitope recognized by an anti-gp14 monoclonal antibody. Materials and methods: Purified EHV-1 DNA (Kentucky strain) was supplied by Rhone-Merieux (Lyon, France). Anti-gp14 monoclonal antibody and phage λ gt11 recombinant 4a1 expressing gp14 epitope were obtained as a gift from Dr. George Allen (University of Kentucky, Lexington, KY). SequenaseTM and T4 gene 32-encoded protein were purchased from United States Biochemical Corp. and Boehringer-Mannheim, respectively. Allen and Yeagan (1987) reported that the *gp14* gene spans the junction between the *Bam*HI restriction fragments *a* and *i* of EHV-1 DNA. The identity of the EHV-1 genomic DNA from Rhone-Merieux and that in the present publication was confirmed by Southern-blot analysis with ³²P-labeled DNA fragments from λ gt11 recombinant 4a1, following the procedure previously described (Guo et al., 1987). The EHV-1 DNA fragments *Bam*HI-*a* (about 21.3 kb) and *i* (about 7.1 kb) (Whalley et al., 1981) were isolated from EHV-1 genome by 0.8% agarose-gel electrophoresis. Plasmid pUC (*Bam*HI-*i*) was constructed by inserting the *Bam*HI-*i* fragment into plasmid pUC8 at the *Bam*HI site. The *Bam*HI-*a* fragment was digested with *Eco*RI and ligated into pUC8 digested with *Eco*RI + *Bam*HI. Since there are two *Eco*RI sites in the *Bam*HI-*a* fragment, only the two fragments at both ends could be ligated with the *Bam*HI + *Eco*RI linearized plasmid. Plasmid pUC (*Bam*HI-*a*-*Eco*RI) harboring a 10-kb insert was selected based on the

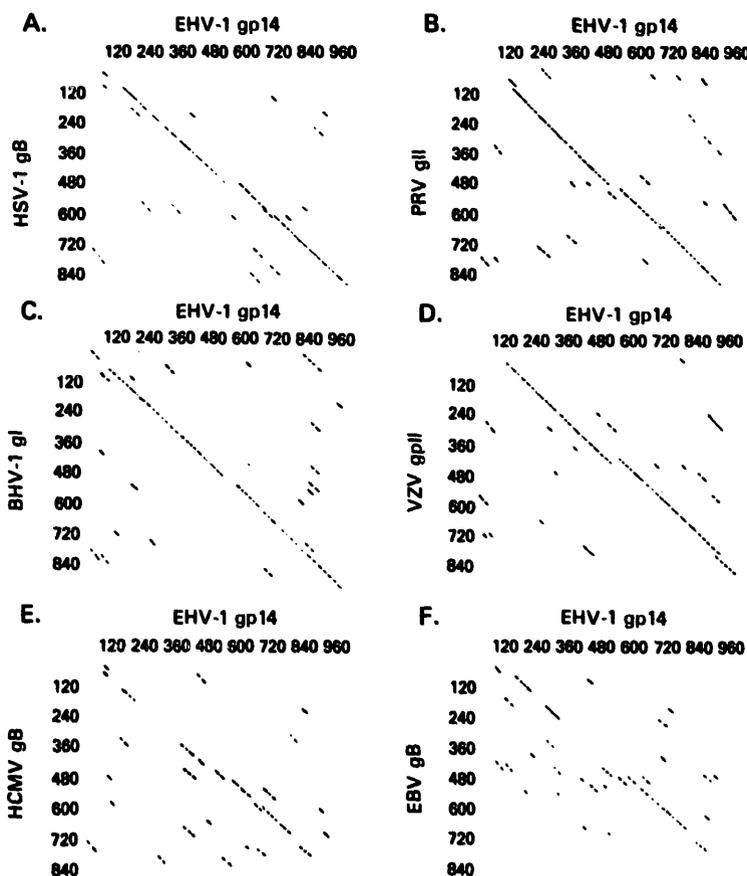


Fig. 2. Homology search of EHV-1 gp14 aa sequence vs. the gB-equivalent glycoproteins of other herpesvirus by DD-matrix with the subprogram of Forward Homology Search from the IBI/Pustell DNA and Protein Sequence Analysis System (IBM computer software, International Biotechnologies, Inc.). Numbers indicate aa residues. The parameters are: range = 10; scale = 0.90; hash level = 1; jump level = 2; step = 2; minimal value plotted = 30. For clarity, the lower-case letters in the original figures were replaced with the dashes (45° angle) when redrawn.

size predicted from the restriction map (Whalley et al., 1981). To facilitate the DNA sequencing, different subclones from pUC (*Bam*HI-*a-Eco*RI) and pUC (*Bam*HI-*i*) were prepared. Nucleotide sequences from the subclones were transferred to the IBM computer files. Regions of sequence overlap were identified and the final sequence was assembled with the IBI/Pustell Sequence Analysis Program. The DNA was sequenced on both strands. Sequencing of the plasmid pUC (*Bam*HI-*a-Eco*RI) was started from the *Bam*HI site, since *gp14* spans the *Bam*HI junction (Allen and Yeagan, 1987). The orientation of the plasmid pUC (*Bam*HI-*i*) was determined by restriction enzyme digestion, the *Bam*HI end closer to the *Eco*RI site was demonstrated to be at the junction of *Bam*HI-*a* and *i*. Sequencing was started from this end. DNA sequencing was performed with Sequenase™ (Tabor and Richardson, 1987) by the dideoxy chain-termination method (Sanger et al., 1977), using double-stranded plasmid templates and the alkaline-denaturation method. The T4 gene 32-encoded protein was utilized to improve sequencing efficiency. In most cases, M13 universal primers were utilized for sequences extending from the termini of each fragment in the subclones. Additional synthetic DNA primers were used to determine sequences further from the fragment termini: These primers were synthesized on a Biosearch 8700 or an Applied Biosystems 380B DNA synthesizer. The DNA fragment corresponding to the gp14 epitope contained in the λ gt11 recombinant 4a1 was isolated from purified λ gt11 DNA after *Eco*RI digestion, following the principle for purifying bacteriophage ϕ 29 DNA (Guo et al., 1986). This fragment was cloned into the *Eco*RI site of pUC8. The nt sequencing was carried out as described above.

stimulate the class-II CTL (Zarling et al., 1986; Martin et al., 1987). gG is a target for antibody directed complement-dependent virus neutralization (Sullivan and Smith, 1988).

It has been demonstrated previously that gp13 of EHV-1 shows some homology to gC of HSV, gIII of PRV and gpV of VZV (Allen and Coogle, 1988; Guo et al., 1989), and that immunization of guinea pigs with recombinant vaccinia virus expressing the gp13-encoding gene protects the host against lethal challenge of EHV-1 (Guo et al., 1989). In this paper, I report the nt sequence of the gp14-encoding gene, compare its aa sequence with those of HSV gB (Bzick et al., 1984), PRV gII (Robbins et al., 1987), BHV-1 gI (Whitbeck et al., 1988), EBV gB (Pellett et al., 1985), VZV gII (Davidson et al., 1986) and HCMV gB (Cranage et al., 1986) and study an antigenic determinant of gp14. Our subsequent studies, truncation and expression of *gp14* in recombinant vaccinia virus and induction of protective immunity in immunized animals (Guo et al., 1990), further confirm that the reported ORF was indeed the gp14-coding sequence.

```

EHV-1 gp14 137- F Y T C P P P T G S T V V R I E P P R T C P K F D L G R - N F T E G I A V -172
                *                               *
PRV gII      123- F Y V C P P P S G S T V V R L E P E Q A C P E Y S Q G R - N F T E G I A V -158
                *                               *
BHV-1 gI     125- F F V C P P P S G A T V V R L A P A R P C P E Y G L G R - N Y T E G I G V -160
                *                               *
VZV gpII     56- F Y V C P P P T G S T I V R L E P T R T C P D Y H L G K - N F T E G I A V - 91
                *                               *
HSV-1 gB     112- F Y V C P P P T G A T V V Q F E Q P R R C P T R P E G Q - N Y T E G I A V -147
                *                               *
HCMV gB      91- Y R V C S M A Q G T D L I R F E R N I I C T S M K P I N E D L D E G I M V -127
                *                               *
EBV gB       48- F R V C E L S S H G D L F R F S S D I Q C P S F G T - R E N H T E G L L M - 84
                *                               *

```

Fig. 3. Sequence alignment of an epitope in gp14 of EHV-1 with the sequences in gB-equivalent gp of other herpesviruses. The identical aa are underlined. Two highly conserved Cys residues were marked with asterisks.

RESULTS AND DISCUSSION

(a) Nucleotide sequence analysis of the *gp14* gene

Analysis of the nt sequence revealed one ORF extending from nt 300 to 3239 (Fig. 1), reading from left-to-right on EHV-1 genome. The ATG start codon was contained in the *Bam*HI-*a*-*Eco*RI fragment, the stop codon TAA in the *Bam*HI-*i* fragment (Whalley et al., 1981; Allen and Yeargan, 1987). Putative transcriptional regulatory signals were found in the region 5' to the ATG codon at nt 300. A TATA box having the sequence AATATAT (nt 154–160) and a potential CAT box, GGTC AATT (nt 75–82), were detected.

A scanning model has been proposed to explain the mechanism by which eukaryotic mRNAs initiate translation (Kozak, 1983; 1986). The cardinal rule of this model requires that ribosomes access initiation by first binding to the 5' end, followed by linear scanning of the mRNA; commitment to initiation is usually at the first 5'-proximal AUG codon, although there are some exceptions (Kozak, 1986; Pelletier and Sonenberg, 1988). A purine at nt position -3 (upstream from the ATG codon) is essential for the initiation of translation and this translation is stimulated by C at nt positions -1 and -2 when the rest of the sequence is suboptimal (Kozak, 1986). The sequence context around the proposed start codon of gp14 was GTCATGT (nt 297–303) which qualified as a functional sequence context for translation initiation of eukaryotic mRNA. This sequence was identical to the corresponding sequence GAGATGT in the *gp13* gene (Allen and Coogle, 1988; Guo et al., 1989) in terms of the -3 and +4 position.

Out of 11 nt in the sequence, 5'-TCCTGCGCGCA-3' (nt 221–231), 9 are complementary to the 18S ribosomal RNA sequence 3'-AGGAAGGCGT-5' (Hagenbuchle et al., 1978). This sequence likely serves as the ribosome-binding site.

A polyadenylation signal, AATAAA (nt 3247–3252), was located 8 nt downstream from the stop codon TAA (nt 3237–3239).

(b) Analysis of the gp14 protein structure

The aa sequence predicted from the ORF codes for a 979-aa protein (about 108 kDa). Since gp14 is a membrane-

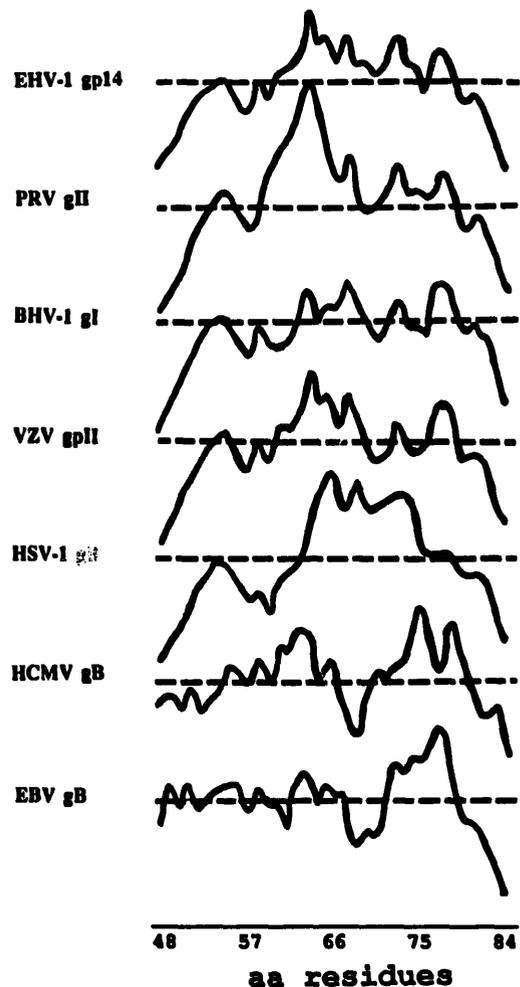


Fig. 4. Comparison of hydrophobic profiles of the epitopes shown in Fig. 3 (Kyte and Doolittle, 1982). Hydrophilic regions were plotted above the x axes (dashed lines). The scale has been compressed to allow multiple profiles on the panel and each profile is plotted one hydrophilicity unit below the previous one. Residue number represents the EBV gB epitope. For other aa see Fig. 3.

bound gp, it is predicted to possess a Sig and a membrane anchor region. Analysis of the aa sequence revealed a number of features common to membrane-bound gp.

(i) *The signal sequence (Sig) domain*

The region extending from aa 67–86 displayed hydrophobic character and is likely the Sig core (Fig. 1). Another common feature of Sig is that the hydrophobic sequence is preceded by a couple of positively charged aa (Perlman and Halvorson, 1983). However, in gp14 the hydrophobic Sig was preceded by an exceptionally long (66 aa) sequence with a preponderance of hydrophilic aa. This unusual feature has been detected in gII of PRV (Robbins et al., 1987) and gI of BHV-1 (Whitbeck et al., 1988) as well.

(ii) *The membrane anchor region and the cytoplasmic domain*

There was a hydrophobic peak extending from aa 831–871 (Fig. 1). The hydrophobic region consisting of 41 aa was assumed to be the transmembrane domain. The hydrophilic cytoplasmic domain was 108 aa in length (Fig. 1).

(iii) *The predicted N-linked glycosylation sites*

There were eleven N-X-T/S sites (where X can be any aa except proline) with potential for N-linked glycosylation (Montreuil, 1984). Interestingly, there were two sites in the cytoplasmic domain (Fig. 1). There was no direct evidence to prove that these sites are glycosylated.

(c) *gp14 is homologous to gII of PRV, gI of BHV-1, gII of VZV, and gB of HSV, EBV and HCMV*

A comparison of the aa sequence of EHV-1 gp14 with the gp of other herpesviruses revealed that gp14 is homologous to PRV gII (Robbins et al., 1987), BHV-1 gI (Whitbeck et al., 1988), VZV gII (Davidson and Scott, 1986), and gB of HSV-1 (Bzick et al., 1984), EBV (Pellett et al., 1985) and HCMV (Cranage et al., 1986) (Fig. 2).

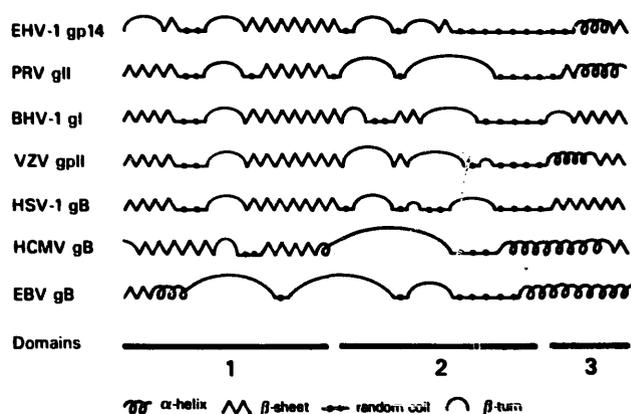
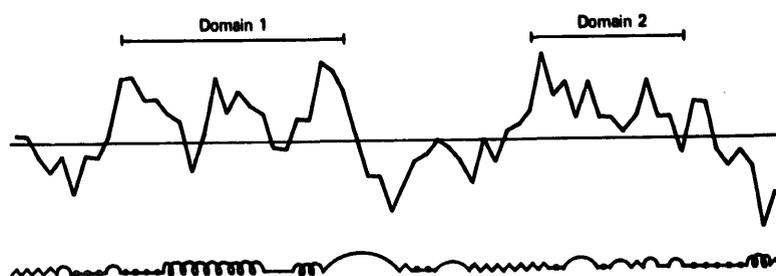


Fig. 5. Secondary structure prediction of the epitopes shown in Figs. 3 and 4. Except EBV, the structures of domains 1 and 3 are very similar in six herpesviruses. Domain 2, containing a preponderance of charged and polar aa (Fig. 4) predicted to be present on the surface of the protein, are more diverse.

(d) *Analysis of an antigenic determinant recognized by an anti-gp14 monoclonal antibody*

Antigenic determinants or epitopes are the portions on the surface of a protein that interact with the specific antibodies. They may be continuous (comprising a single segment of peptide chain) or discontinuous (comprising two or more chain segments brought together by the folding of the polypeptide chain) (Hopp, 1986; Westhof et al., 1984; Berzofsky et al., 1985).

The λ gt11 expression system represents an ideal approach to identify the epitopes for specific monoclonal antibodies (Mehra et al., 1986). Bacteriophage λ gt11 recombinant 4a1 expressing a gp14 epitope recognized by anti-gp14 monoclonal antibody 3F6 was constructed by Allen and Yeargan (1987). In their previous work, a library of the EHV-1 genome was constructed in the λ gt11 expression vector. Recombinant bacteriophages expressing EHV-1 gp14 epitopes as fusion products with β Gal were detected by immuno-screening with anti-gp14 monoclonal



107-TVDPTLLPTETPDPLRLAVRESGILAEDGDFYTCPPPTGSTVVRIEPPRTCPKFDLGRNFTEGIIV-172

Fig. 6. Hydrophobic plot and secondary structure prediction of aa sequence coded by the EHV-1 DNA fragment in gt11 recombinant 4a1. Hydrophilic regions were plotted above the x-axes. See Fig. 5 for symbols.

antibodies, which was identified previously by Western-blot analysis and verified by immunoprecipitation with [³H]glucosamine-labeled gp14. Phage λ gt11 4a1 is one of the recombinant λ bacteriophages expressing gp14 epitopes that are recognized by the monoclonal antibody specific for gp14. The EHV-1 DNA fragment contained in this recombinant bacteriophage has been used as a hybridization probe for mapping the *gp14* gene (Allen and Yeargan, 1987). The viral DNA fragment from the λ gt11 recombinant 4a1 was recloned and sequenced. Sequence comparison indicated that this fragment, coding for a 66-aa polypeptide (aa 107–172), originated from the N terminus of the deduced surface domain of gp14 (Fig. 1). Sequence alignment revealed that aa residues 137–172 were the best conserved regions among the gB glycoproteins of the EHV-1, PRV, BHV-1, VZV, HSV-1, HCMV and EBV viruses (Figs. 2–5).

One feature of antigenic sites is that they usually contain a large number of charged and polar aa. Hydrophilicity analysis and secondary structure prediction have been utilized to identify antigenic determinants. Typical contiguous antigenic determinants are hydrophilic stretches of aa adjacent to regions having potential for secondary structures involving β -turns. Hydrophobic plots and protein secondary-structure predictions indicate that there are two domains; aa residues 116–134 (Fig. 6, domain 1) and 152–166 (Fig. 6, domain 2) possess the features of typical antigenic determinants. Homologous comparison with the gB-equivalent gp of other herpesviruses indicates that the hydrophilic stretches are the less conserved region.

(e) Conclusions

Nucleotide sequence analysis has revealed an ORF of 2940 bp. An analysis of the deduced aa sequence indicated that an antigenic determinant for a monoclonal anti-gp14 antibody was present in the N terminus of the protein. Alignment of aa sequences indicates strong homology to the gB-equivalent gp of other herpesviruses. Since no intron has been discovered in herpesvirus genes, these data suggest that the ORF codes for gp14 which was the counterpart of herpesvirus gB.

ACKNOWLEDGEMENTS

I thank Mr. Hector L. Martinez and Dr. Gerard Johnson who synthesized the oligodeoxyribonucleotides, and Drs. Enzo Paoletti, Charles de Taisne, Marion Perkus, Randall Weinberg and Gerard Johnson for helpful discussions. The work was supported by Virogenetic, Inc. I appreciate Dr. George Allen for the communication of his unpublished results as well as Dr. Trevor Prior, Dr. Michael Bodenterch,

Mr. Vel Karacostas and Mrs. Betty Morris for their comments on the manuscript.

REFERENCES

- Allen, G. and Coogle, L.D.: Characterization of an equine herpesvirus type 1 gene encoding a glycoprotein (gp13) with homology to herpes simplex virus glycoprotein C. *J. Virol.* 62 (1988) 2850–2858.
- Allen, G.P. and Yeargan, M.R.: Use of *lambda* gt11 and monoclonal antibodies to map the genes for the six major glycoproteins of equine herpesvirus 1. *J. Virol.* 61 (1987) 2454–2461.
- Berzofsky, J.A.: Intrinsic and extrinsic factors in protein antigenic structure. *Science* 229 (1985) 932–940.
- Bzick, D.J., Fox, B.A., DeLuca, N.A. and Person, S.: Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex virus type 1. *Virology* 133 (1984) 301–307.
- Cantin, E.M., Eberle, R., Baldick, J.L., Moss, B., Willey, D.E., Nots, A.L. and Openshaw, H.: Expression of herpes simplex virus 1 glycoprotein B by a recombinant vaccinia virus and protection of mice against lethal herpes simplex virus 1 infection. *Proc. Natl. Acad. Sci. USA* 84 (1987) 5908–5912.
- Cranage, M.P., Kouzarides, T., Bankier, A.T., Satchwell, S., Weston, K., Tomlinson, P., Barrell, B., Hart, H., Bell, S.E., Minson, A.C. and Smith, G.L.: Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. *EMBO J.* 5 (1986) 3057–3063.
- Davidson, A.J. and Scott, J.E.: The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* 67 (1986) 1759–1816.
- Glorioso, J., Kees, U., Kumel, G., Kirchner, H. and Krammer, P.: Identification of herpes simplex virus type 1 (HSV-1) glycoprotein gC as the immunodominant antigen for HSV-1-specific memory cytotoxic T lymphocytes. *J. Immunol.* 135 (1985) 575–582.
- Guo, P., Grimes, S. and Anderson, D.: A defined system for in vitro packaging of DNA-gp3 of the *Bacillus subtilis* bacteriophage ϕ 29. *Proc. Natl. Acad. Sci. USA* 83 (1986) 3505–3509.
- Guo, P., Erickson, S. and Anderson, D.: A small viral RNA is required for in vitro packaging of bacteriophage ϕ 29 DNA. *Science* 236 (1987) 690–694.
- Guo, P., Goebel, S., Davis, S., Perkus, M.E., Languet, B., Desmettre, P., Allen, G. and Paoletti, E.: Expression of equine herpesvirus type 1 gene coding for glycoprotein 13 in recombinant vaccinia virus and protection of immunized animals. *J. Virol.* 63 (1989) 4189–4198.
- Guo, P., Goebel, S., Perkus, M.E., Taylor, J., Norton, E., Allen, G., Languet, B., Desmettre, P. and Paoletti, E.: Coexpression by vaccinia virus recombinants of equine herpesvirus-1 glycoproteins gp13 and gp14 results in potentiated immunity. *J. Virol.* 64 (1990) (In press).
- Hagenbuchle, O., Santer, M., Steitz, J.A. and Mans, R.J.: Conservation of the primary structure at the 3' end of 18S rRNA from eukaryotic cells. *Cell* 13 (1978) 551–563.
- Hopp, T.P.: Protein surface analysis. *J. Immunol. Meth.* 88 (1986) 1–18.
- Kozak, M.: Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* 47 (1983) 1–45.
- Kozak, M.: Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44 (1986) 283–292.
- Kyte, J. and Doolittle, R.F.: A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157 (1982) 105–132.
- Martin, S., Moss, B., Berman, P.W., Laskey, L.A. and Rouse, B.T.: Mechanisms of antiviral immunity induced by a vaccinia virus recombinant expressing herpes simplex virus type 1 glycoprotein D: cytotoxic T cells. *J. Virol.* 61 (1987) 726–734.

- Mehra, V., Sweetser, D. and Young, R.A.: Efficient mapping of protein antigenic determinants. *Proc. Natl. Acad. Sci. USA* 83 (1986) 7013-7017.
- Montreuil, J.: Spatial conformation of glycans and glycoproteins. *Biol. Cell* 51 (1984) 115-132.
- Perlman, D. and Halvorson, H.O.: A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* 167 (1983) 391-409.
- O'Callaghan, D.J., Henry, B.E., Wharton, J.H., Dauenhauer, S.A., Vance, R.B., Staczek, J., Atherton, S.S. and Robinson, R.A.: Equine herpesviruses: biochemical studies on genomic structure, DI particles, oncogenic transformation, and persistent infection. In Becker, Y. (Ed.), *Developments in Molecular Virology*, Vol. 1, Herpesvirus DNA. Nijhoff, The Hague, 1981, pp. 387-418.
- Paoletti, E., Lipinkas, B.R., Samsonoff, C., Mercer, S. and Panicali, D.: Construction of live vaccines using genetically engineered poxvirus: biological activity of vaccinia virus recombinants expressing the hepatitis B virus surface antigen and the herpes simplex virus glycoprotein D. *Proc. Natl. Acad. Sci. USA* 81 (1984) 193-197.
- Pelletier, J. and Sonenberg, N.: Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334 (1988) 320-325.
- Pellett, P.E., Biggin, M.D., Barrell, B.L. and Roizman, B.: Epstein-Barr virus genome may encode a protein showing significant amino acid and predicted secondary structure homology with glycoprotein B of herpes simplex virus 1. *J. Virol.* 56 (1985) 807-813.
- Robbins, A.K., Dorney, D.J., Wathen, M.W., Whealy, M.E., Gold, C., Watson, R.J., Holland, L.E., Weed, S.D., Levine, M., Glorioso, J.C. and Enquist, L.W.: The pseudorabies virus gII gene is closely related to the gB glycoprotein gene of herpes simplex virus. *J. Virol.* 61 (1987) 2691-2701.
- Rosenthal, K.L., Smiley, J.R., South, S. and Johnson, D.C.: Cells expressing herpes simplex virus glycoprotein gC but not gB, gD, or gE are recognized by murine virus-specific cytotoxic T lymphocytes. *J. Virol.* 61 (1987) 2438-2447.
- Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74 (1977) 5463-5467.
- Sullivan, V. and Smith, G.L.: The herpes simplex virus type 1 US7 gene product is a 66K glycoprotein and is a target for complement-dependent virus neutralization. *J. Gen. Virol.* 69 (1988) 859-867.
- Tabor, S. and Richardson, C.C.: DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* 84 (1987) 4767-4771.
- Westhof, E., Altschuh, A., Moras, D., Bloomer, A.C., Mondragon, A., Klug, A. and Regenmortel, M.: Correlation between segmental mobility and the location of antigenic determinants in protein. *Nature* 311 (1984) 123-126.
- Whalley, J.M., Robertson, G.R. and Davison, A.J.: Analysis of the genome of equine herpesvirus type 1: arrangement of cleavage sites for restriction endonucleases *EcoRI*, *BglII*, and *BamHI*. *J. Gen. Virol.* 57 (1981) 307-323.
- Whitbeck, J.C., Bello, L.J. and Lawrence, W.C.: Comparison of the bovine herpesvirus 1 gI gene and the herpes simplex virus type 1 gB gene. *J. Virol.* 62 (1988) 3319-3327.
- Zarling, J.M., Moran, P.A., Lasky, L.A. and Moss, B.: Herpes simplex virus (HSV)-specific human T-cell clones recognize HSV glycoprotein D expressed by a recombinant vaccinia virus. *J. Virol.* 59 (1986) 506-509.