Antigenic and Protein Sequence Homology between VP13/14, a Herpes Simplex Virus Type 1 Tegument Protein, and gp10, a Glycoprotein of Equine Herpesvirus 1 and 4

GARY R. WHITTAKER,¹ MARCELLO P. RIGGIO,² IAN W. HALLIBURTON,¹ RICHARD A. KILLINGTON,¹
GEORGE P. ALLEN,³ AND DAVID M. MEREDITH^{1*}

Department of Microbiology, University of Leeds, Leeds LS2 9JT,¹ and Department of Veterinary Pathology, University of Glasgow, Glasgow G61 1QH,² United Kingdom, and Department of Veterinary Science, University of Kentucky, Lexington, Kentucky 40546³

Received 8 November 1990/Accepted 30 January 1991

Monospecific polyclonal antisera raised against VP13/14, a major tegument protein of herpes simplex virus type 1 cross-reacted with structural equine herpesvirus 1 and 4 proteins of M_r 120,000 and 123,000, respectively; these proteins are identical in molecular weight to the corresponding glycoprotein 10 (gp10) of each virus. Using a combination of immune precipitation and Western immunoblotting techniques, we confirmed that anti-VP13/14 and a monoclonal antibody to gp10 reacted with the same protein. Sequence analysis of a λ gt11 insert of equine herpesvirus 1 gp10 identified an open reading frame in equine herpesvirus 4 with which it showed strong homology; this open reading frame also shared homology with gene UL47 of herpes simplex virus type 1 and gene 11 of varicella-zoster virus. This showed that, in addition to immunological cross-reactivity, VP13/14 and gp10 have protein sequence homology; it also allowed identification of VP13/14 as the gene product of UL47.

The tegument region of the herpes simplex virus (HSV) particle was originally defined as an electron-opaque, amorphous region located between the virus envelope and capsid (27); the proteins composing this region were defined by a subtractive method as those not released by nonionic detergent and not present in the virus capsid. The tegument contains five major proteins, VP1/2, VP13, VP14, VP16, and VP22, each of which constitutes $\geq 5\%$ of the mass of the virus particle (17, 27).

VP13 and VP14 have been considered to be related through some form of posttranslational modification (17) and are thought to be the product of a single gene (21a). The two polypeptides exhibit polymorphism in molecular weight between HSV type 1 (HSV-1) strains; variation in apparent molecular weight of one polypeptide is mimicked by variation in the molecular weight of the other (23). All isolates of HSV-2 appear to contain a single invariant polypeptide designated VP14 (5). VP13 and VP14 of HSV-1 have been shown to be phosphorylated within the virion; the gene coding for these polypeptides has been mapped to a region between map positions 0.66 and 0.76 on the HSV-1 genome (18).

The studies reported here were carried out as part of a series of experiments attempting to characterize the tegument proteins of HSV-1, to identify the genes coding for them, and to identify homologous proteins in other herpesviruses.

MATERIALS AND METHODS

Cell culture, virus strains, and virus purification. HSV-1 strain SC16 and HSV-2 strain 3345 were cultivated in

BHK-21 cells grown in the autoclavable form of Eagle medium (Flow Laboratories) containing 10% calf serum and 10% tryptose phosphate broth. Equine herpesvirus 1 (EHV-1) strain Ab-1 and EHV-4 strain MD were cultivated in either RK13 cells (EHV-1) or NBL-6 cells (EHV-4) as described previously (22). HSV-1, EHV-1, and EHV-4 were purified from extracellular virus by standard techniques (22), and HSV-2 was purified from cytoplasmic extracts of cells (27).

SDS-PAGE and Western immunoblotting. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting were performed essentially as described previously (22) with gels containing 9 or 12% acrylamide cross-linked with *N,N'*-diallyltartardiamide. Western blots were developed with Vectastain ABC kits and peroxidase-linked antibody (Vector Labs), and color was developed with chloronaphthol- H_2O_2 .

Immune precipitation. For immune precipitation of glycoprotein 10 (gp10), RK cells were infected with EHV-1 at 1 PFU/cell and labeled at 6 h postinfection with 10 μ Ci of [35 S]methionine (Amersham) per ml in medium containing 20% of the normal levels of methionine. Cells were harvested at 24 h postinfection, solubilized in RIPA buffer (0.1% sodium dodecyl sulfate, 0.5% Nonidet P-40, 0.1 M NaCl, 50 mM Tris-HCl [pH 7.4]), and centrifuged at $100,000 \times g$ for 1 h; the supernatant was removed and used as an antigen source. Monoclonal antibody was incubated with a protein A-Sepharose suspension (Pharmacia) for 1 h at room temperature, washed three times with phosphate-buffered saline, and incubated with antigen at 4°C overnight. The antibody was then washed three times with RIPA buffer and then with SDS-PAGE solubilizing buffer at 100°C for 5 min before electrophoresis and Western blotting.

* Corresponding author.

FIG. 1. Western blot of a 9% acrylamide gel. (a) Reactivity of R220 (anti-HSV-1 VP13/14) with HSV-1 (lane 1), HSV-2 (lane 2), EHV-1 (lane 3), and EHV-4 (lane 4). (b) Reactivity of 13A9 (anti-EHV-1 gp10) with EHV-1 (lane 1) and EHV-4 (lane 2). Molecular weights of reactive species are indicated in thousands.

Antisera. Monoclonal antibody 13A9 (2) reacts with gp10 of EHV-1. R220 is a monospecific polyclonal antiserum raised in rabbits against SDS-PAGE-purified HSV-1 VP14.

Plasmids and cloning of EHV-4 DNA. The plasmid containing the 9.6-kb *Bam*HI G fragment of EHV-4 (strain 1942) cloned in pUC9 (6) was digested with *Sal*I, and the resulting left-terminal 1.9-kb *Bam*HI-*Sal*I fragment and the adjoining 0.9-kb *Sal*I subfragment of *Bam*HI-G were separately cloned into the *Sal*I site of the plasmid vector pUC8 (29) to generate plasmids pUC1.9G and pUC0.9G, respectively.

Bacterial strains. *Escherichia coli* JM101 was used for all recombinant DNA experiments. Recombinant plasmids were selected on plates containing 100 µg of ampicillin per ml.

Isolation of plasmid DNA. Plasmid DNA was isolated from bacteria by the boiling method (15) for small-scale miniprep- aration analysis. Large-scale isolation of plasmid DNA from bacteria was carried out by the alkaline lysis method (21); the DNA was further purified by banding on cesium chloride gradients containing 50% (wt/vol) cesium chloride and 200 µg of ethidium bromide per ml. Plasmid DNA was harvested, and ethidium bromide was removed by multiple extractions with isopropanol. The DNA was dialyzed against 0.1× TE (1 mM Tris-HCl, 0.1 mM EDTA [pH 8.0]) and recovered by ethanol precipitation.

DNA sequencing. The 1.9 kb *Bam*HI-*Sal*I and 0.9-kb *Sal*I subfragments of *Bam*HI-G contained in plasmids pUC1.9G and pUC0.9G, respectively, were sequenced as previously described (25) by using the dideoxy chain termination method (26) and [α - 35 S]dATP as a label (4). The Klenow fragment of *E. coli* DNA polymerase I was used for chain elongation at 37°C; the temperature was increased to 45°C when regions of DNA prone to secondary structure formation were sequenced. Plasmid DNA was sequenced by using pUC-specific sequencing primers, and EHV-4-specific synthetic deoxyoligonucleotide primers were used to generate

FIG. 2. (a) Autoradiography of a 12% acrylamide gel to show immune precipitation of [35 S]methionine-labeled gp10 by 13A9. (b) Western blot of material immune precipitated by 13A9 and probed with 13A9 (lane 1) and R220 (lane 2). Molecular weights of reactive species are indicated in thousands. Also shown is reactivity of immunoglobulin released from the immune precipitation matrix (◀).

sequence data rapidly. Each nucleotide was sequenced at least three times until any ambiguities were totally removed. Sequence data were compiled and analyzed with the Beckman Microgenie sequencing programs.

Nucleotide sequence accession number. The primary nucleotide sequence of open reading frame B6 has been submitted to the EMBL data library under accession number X 17684 B6.

RESULTS

Cross-reactivity of polyclonal antisera specific for HSV-1 tegument proteins. As part of the characterization of the tegument proteins of HSV-1, polyclonal antisera were raised against individual polypeptides and tested for cross-reactivity against a range of α -herpesviruses. One antibody in particular, R220, showed cross-reactivity. R220 was raised against purified VP14 of HSV-1, and the reaction of this antibody with HSV-1, HSV-2, EHV-1, and EHV-4 on a

λ gt11/EHV-1 insert

Q A P R P R A D F A P P S G E E S S	18
GAATTCCAAGCACCACGACCTCGCGCGGACTTTGCACCGCCCTCCGGCGAGGAATCATCT	60
EcoRI	
S E E E E E E G P A Q A P L D E E D Q L	38
AGCGAGGAAGAGGAGGAAGAGGGTCCGCGCAAGCTCCGCTGGACGAGGAAGACGACTA	120
M Y A D Q Y S V G D S S D E N D E E E D	58
ATGTATGCTGACCACTACTCTGTAGGGGACTCTAGTGACGAAACGACGAGGAAGAAGAC	180
P R L G S D Y P T S A E S	71
CCCCGTCTAGGATCTGACTATCCACGTCGCGCAATCCAGAATTC	226
EcoRI	

FIG. 3. Nucleotide and predicted amino acid sequence of the λ gt11 insert containing the epitope recognized by monoclonal antibody 13A9. The *Eco*RI sites are indicated.

EHV-4 B6
EHV-1 insert

98	Q	A	P	R	S	R	A	D	F	A	P	P	P	E	E	D	S	S	S	E
1	Q	A	P	R	P	R	A	D	F	A	P	P	S	G	E	E	S	S	S	E
118	E	E	D	E	E	G	P	S	Q	A	P	L	D	E	E	D	Q	L	M	Y
21	E	E	E	E	E	G	P	A	Q	A	P	L	D	E	E	D	Q	L	M	Y
138	A	D	Q	Y	S	V	G	N	S	S	D	D	N	E	E	D	Y	L	Q	P
41	A	D	Q	Y	S	V	G	D	S	S	D	E	N	D	E	E	D	P	R	
158	E	V	E	-	Y	P	T	S	A	E	S									
61	L	G	S	D	Y	P	T	S	A	E	S									

FIG. 4. Amino acid comparison of the EHV-1 λ gt11 insert reacting with 13A9 and the homologous region of open reading frame B6 of EHV-4.

Western blot is shown in Fig. 1a. For HSV-1 a characteristic doublet of 84,000 and 79,000 M_r , corresponding to VP13 and VP14 of HSV-1, was seen, whereas for HSV-2 only a single protein of 89,000 M_r , corresponding to VP14, was detectable. Of the other herpesviruses tested only EHV-1 and EHV-4 were seen to consistently react with R220; for each virus a single band was visible (120,000 M_r for EHV-1, 123,000 M_r for EHV-4).

The molecular weights of the EHV-1 and EHV-4 polypeptides that reacted with R220 were very similar to those reported for gp10 (1), which raised the possibility that gp10 was reacting with R220. A monoclonal antibody to EHV-1 gp10, termed 13A9, has been shown to react with a [3 H]glucosamine-labeled polypeptide of around 120,000 M_r (2). The reactivity of 13A9 on a Western blot is shown in Fig. 1b. This antibody reacted with polypeptides of 120,000 M_r in EHV-1 and 123,000 M_r in EHV-4, but no reactivity could be observed for 13A9 when tested against either HSV-1 or HSV-2 (data not shown).

To determine whether R220 and 13A9 were reacting with the same protein, gp10 was first immune precipitated from EHV-1-infected cells (Fig. 2a). This material was then eluted from the immune precipitation matrix, subjected to SDS-

PAGE and Western blotting, and then probed with R220. R220 reacted with same polypeptide immune precipitated by 13A9 (Fig. 2b).

Identification of the open reading frame encoding EHV-4 gp10. A λ gt11 expression system has been used to map the gene encoding EHV-1 gp10 to a map positions 0.093 to 0.114 on the EHV-1 genome (2). To identify the gene coding for gp10, the λ gt11 insert containing the epitope that reacts with 13A9 was sequenced (Fig. 3). No significant homology could be found when this sequence was compared with the complete DNA sequence of HSV-1 (19). However, sequence data were available for a region of the EHV-4 genome between 0.067 and 0.122 map units (M. Riggio, unpublished data); when the EHV-1 λ gt11 insert was compared with this, a region of strong homology was found. A comparison of the predicted amino acid sequences of this region is shown in Fig. 4.

The homologous region in EHV-4 is part (residues 96 through 169) of an open reading frame designated B6 (Fig. 5), which encodes a gene product of 872 amino acids that is predicted to have a molecular mass of 97.4 kDa. The DNA and predicted amino acid sequence of EHV-4 B6 are shown in Fig. 6, and a plot of the hydropathic profile of the polypeptide is shown in Fig. 7. Features of the B6 gene product include a large stretch of around 200 amino acids toward the N terminus which is both hydrophilic and highly acidic; the region between residues 111 and 201 is especially charged, containing 26 glutamic acid residues. There appear to be no hydrophobic regions of the polypeptide characteristic of either a signal sequence (24) or a membrane-spanning domain (10).

Homology of B6 with HSV-1 and VZV. When open reading frame B6 was compared with the DNA sequences of HSV-1 (19) and varicella-zoster virus (VZV) (8), the gene product of B6 showed 17% identity with the HSV-1 UL47 gene product and 27% identity with the VZV gene 11 product. A comparison of the predicted products of these three genes is shown in Fig. 8. The hydrophilic N terminus appears to be a characteristic of these genes; it is conserved in all three but with noticeable differences. EHV-4 and VZV both possess a glutamic acid-rich region; instead, HSV-1 contains an arginine-rich region, giving this protein a highly basic N terminus.

FIG. 5. Map of the EHV-4 genome showing the restriction enzyme sites and DNA fragments used in studying the genome. The genome consists of a unique long region and a unique short region flanked by inverted repeats. (a) Location of the BamHI restriction sites in the viral genome. (b) Genomic location of the BamHI G fragment and the 1.9-kb BamHI-SalI fragment and the 0.9-kb SalI-SalI fragment sequenced during this study. (c) Location of open reading frame B6. The dot represents the 5' end and the arrowhead represents the 3' end of the gene.

GGATCCAGTTCCACCGCTTACCGTAGGGATTTCGTGAGCTGCTGAAACGCTTCTTCCGTCTAACCTCACCCACAGAGCATGGAACTGACGTTCTTGACTACTCATCTATTTCAAG	120
D P V P P L T V G I R Q T A E T L A L P S N L T L Q S M E T D V L D Y S S I S G	
CGACGAGCTCAACCAGATGTTTGACATTTAATAACAATAAGCACGTTTCCAACTTAACATAATGCCGATATTTCCGTGATACGCTGCGTGAATAGAAGTAATGGGGGAGGTGGGC	240
D E L N Q M F D I -	
GTGGTCTGCGGGTGGTGTATGTTAAATGGGCCCCGAGGCTATAGGCAAGTTTGTGTCATTGCTGATCTGCTGCAACAAACGACAATTAACACCAATCTTCAAATATCGCCCAT	360
TAACAGTACAAAAGTGGGGGTATGGCGGTTTTGAAGCTCGTAGCTTGCTATATAAACTCGCGCCCTTTGCCGAGATGGATGTTGCTATCTAGCGTAGATAGCGGGGCTTTGCCGTC	480
AAAAGTACGCGTTGTAAGTACAGCGATACGGAAGTAGTATGACATGGACCAACATCACGCGGTTCCGCGTGGGGCGCTATACGAGGCTCGCAGATCAATAGAAACGCGCTCCCATCCAT	600
(ORF B6) M D Q H H G V R G G A P I R R P R R S I E T R S H P	26
TTAGAGCCGAGGAAATACACAGCGACATACAGCACGCCAAGACTTAGTTATAGAGATGGATTGTCTGGCAGAGCCTTCTCACTGAACCCGGGGCCAAAGCTCACGATCAAAATGAGA	720
F R A A G N T Q R T Y S T P R L S Y R D G L S G R A S S L E P G G Q A H D Q N E	66
GCTCTACACAAAGTACTTCAAATAACCAAGCACCTCATTTTGGGATATCTACGAAGAGTTTTTTCAGATGATGCCCGCGCAGCCACAAGCAAGGCTCGCGCTGATTTTG	840
S S T Q S T S N N Q P S T S F W G Y L R R V F S D D A P A Q P Q A P R S R A D F	106
CTCTCCCCCGAGGAGGACTCATCCAGCGAGGAAGAAGCAGGAAGGTCCTCACAGCTCCGTGGATGAGGAGGACGAGCTCATGTATGCTGACCAATACTCAGTAGGTAAGTCTA	960
A P P P E E D S S S E E E D E E G P S Q A P L D E E D Q L M Y A D Q Y S V G N S	146
GTGATGATAACGAAGACTACCTACAGCCAGAAGTTGAATATCCAATCTCCGAGAATCTGGCGAATATCATAACAGTGGGATGTTGCAGAAGAGGAGCCGGAAGCGAGTCTGAGT	1080
S D D N E E D Y L Q P E V E Y P T S A E S G E Y H N S G M F A E E E P E S E S E	186
CAGACATGGAAATACGAAACGTACGAGGAAATGATACGGAAGTATATCAGATGATAGCCATAGACTTACTCGTACGTGGTGGATAGGTCTATACGCTTAATGGACGACGCACTTG	1200
S D M E N Y E T Y E E N D T E V I S D D S H R L T R T W L D R S I R L M D D A L	226
CACAGTCTTCTGAAATTTCTAAGGCTACTCAATCTACGCGCAGGTTATACGATAGCCAGTTTACTCCAGGGGTCGAGGCTACAAACAAACGGAACCCCTCCAGCGTTTGGTTC	1320
A Q S S E I S K A I T K S T R R L Y D S Q F T P G G R G Y K Q T E T P S Q R L V	266
ATCTATCACGCGCTGGTATGTACGATTCTGACGAAATCGTTATGACAGGGGATTACATGGAGGTGACGACGACCAACAGCGCTTACCAGTATGGGTGCGCGCTATTCAACCCCGG	1440
H L S R A G M Y D S D E I V M T G D Y M E V D D D P N S A Y Q S W V R A I H H P	306
TTGCCATGAACCCATCATGGGAGGAAACAAATTTCCAATCACACCAATACATCTTTCTCGGCACATAGACTATGATATAGACGAGCTAATCGAAATGAATTTGGCGCGAACACCCCGAG	1560
V A M N P S W E E T I S N H T N T S F S D A I D Y D I D E L I E M N L A R T P P	346
TGTTTGGAGGATTGCTAGACAGCGCAGACTTTTTTACAGACTACCCATGCTCTATACATATGCTACTATCACTCAAGACGAGGCTACGAAGAGCGGCGAGCATGGTCTAATACACAGG	1680
V F E G L L D S A D F F Y R L P M L Y T Y A T I T Q D E A Y E E R Q A W S N T Q	386
CGCTGCATGGACACGAACAAAGTTCTTGGCCAGCGCTTGAGTATTACTCTAAGGGGGGATGTACGTGTCCTTACTCAGGAACCCCGGGGATATGGCGACGCGCTAAAACAAG	1800
A L H G H E Q S S W P A L V S D Y S K G G M Y V S P T Q E P R G I W R R A L K Q	426
CAATGGCTCTTCAAGCTAAGCTATGTGTGCTTGGTTAACAGAAATTTGTAACAAAGCGTAGCTCACACAACACCAATTCAAGTGAACCTTTTTTGGTGCAGTCTGCTCTTGAACAGCAA	1920
A M A L Q L K L C V L G L T E F V T K R E L T Q H H S A V T F L V D S L L R T A	466
AAAATTTGTTACTTGGCCAGCGACTTTTAGTATTTGCCTGGGAAAGACGAGGAAAGTGGTGTACGACGCCAGCAGAGCCCTCATAGCACTCTCGGGGTTACGCTCTCCAGCCGC	2040
K N C Y L A S R L L V F A W E R R R E T G V R R P A E P L I A L S G V T L L Q P	506
TTCCCCAGAAAGTCTCAGAAATTAAGTGTAGCAGCGTACATTTGATATAGGGTTGCGCACCCCCAAAGTGGAGTGTTTAGAGCGTCTTTCGACCGCTTGTGTTATGGGCAGAACTACGCA	2160
L P P E V S E L L E Q R T F D I G L R T P Q S G V F R A F F G P L V Y W A E L R	546
GAGCCTTGGCAGACCGAGCTGCCATAAACTGTCGCTATGTTGGATTTCATCTCCAAACATCAGAAATTTATTTATTTGGCAGCGCCCACTCTGCCAGCCCGGGTACACCAAGAAGAAC	2280
R A L R D P A A I N C R Y V G F H L Q T S E I Y L L A R A H S A S P G Y T K E E	586
TGGTTGCAATGGAGGCAACGCTCACACTTGGGACCCCTCATGTTAGAGGTAGCGCTACAGTGGATACAGTGGCCAGTGCACAGTTACTTAGCGAAACGATGCACTGAAAGCTTTTAGGC	2400
L V A M E A T L T L G T L M L E V A L Q W I H V A S A Q L L S E N D A L K A F R	626
GTGTGAGTGCCTATTTCCCAACGCGCTGGCGCCACTTGGTAGCATACGCCTACAGCAGCAGAGTTTGAAGTGCTAAGCAACCCAGATGTGATGGTGGCAGGTGATGAAACCGCCCTGA	2520
R V S A S I P H A L A P L G S I R L H D A E F E V L S N P D V M V A D E T A L	666
GCCAGGCGTTGTTCTTGGAATTTTTCTGTTAGGACCGCACTAAGTGGTGCATGCGTGACTATGCTAATGAGGTGGATGGGGATCTAAGAGACCGTTACTGGTTGTTTTTGGGCG	2640
S Q A L F L G Y F S V R T A L T A C M R D Y A N E V D G G S K E T V T G L F L G	706
TGGGCTAATTAATCAGCGCTCGCTGGCCATATGAACTTTTTACTAACTGTATGGCGGCGCGGCACTTTATGGCGGTAGCAAAATCGCCATACACTCATTAACTCTGCCAGATACA	2760
V G L I I Q R L A G H M N F L L N C M A G A A L Y G G S K I A I H S L T L P R Y	746
GCCTATTGGCGGATGTTATGGCCCTATGCTTCAAGCAGCAGTCTTTGGTGCAGTTTGGCGGCCAGAGACGACATGTTGGAGGAAGTAAAGTAAACACACCGCCCTGGACCCCAACGC	2880
S L L A D V M A P M L Q Q Q S L V D F W R A R D D M L E E L E I T P R P G P P T	786
AAGGCAAGCGGTGGTGGTGGAGATGCCTTTGCCCTCGGACGATCTTCCAGCTATGACTCCAGTGGCCAAAGTAAACAATGGCGCCGTTTGGGGCGCATGGTGACATGGCCAAACACT	3000
Q G K R V V L E M P L P S D D L P A M T P S G Q V N N G A G L G R M V D M A K H	826
TACAGCACTATAGAGAAACAAATATCGGAGACGATGCCTCTTCTGTAGGTAAACGTGGCTTAATGAAATCTGGTGTGGCGGTACGCCATGCGCTGGAGCGGAGAAAGTAATAGAT	3120
L Q H Y R E T I I G D D A S S S V G K R G L M K S G V G V R H A L E A E K V I R	866
ACTCACCCAAAGCACTTAATGCTGTTACGTCCCCGGTATGCTCTACATTCGCAAGCACTTTCATGAAACCTCTTCTACTTACCTAGCACCAACTTGTGTTGACGCTCTCGTAACA	3240
Y S P K S T -	872
ATCTATACATTAAGTGAATACAATGGAAGCTAGTGGGTCTGCTCATGGGCCCGGTTTCCAAAACCTAATCGAGCGCGTGCAGTCAAAGGTCCTCTTGGCAGCCCAAGCGATGT	3360
M E A S G S A S W A R V S K N L I E R R A V K G C L L P T P S D V	33
TATGGACGCTGCTGTTATGGCCTTAAAGACGAACGA	3397
M D A A V M A L K D E R	45

FIG. 6. Nucleotide and predicted amino acid sequence of open reading frame B6.

FIG. 7. Hydrophobicity plot of open reading frame B6 made with the algorithm of Kyte and Doolittle (11) and a moving window of six residues. The hydrophilic region (—glu—) and the glutamic acid-rich region (—glu—) are indicated. Positive hydrophobic indices indicate above-average hydrophobicity, and negative indices indicate above-average hydrophilicity.

DISCUSSION

The results reported herein indicate that the tegument protein VP13/14 and the glycoprotein gp10 of EHV-1 and EHV-4 show both immunological cross-reactivity and protein sequence homology. gp10s of EHV-1 and EHV-4 have been considered major structural glycoproteins, based on incorporation of [^3H]glucosamine (1); EHV-1 gp10 has also been labeled by various in vitro labeling methods with UDP-[^{14}C] galactose and [^3H]borohydride (28). This protein was therefore assumed to be a typical herpesvirus membrane protein; i.e., it contains both a hydrophobic signal sequence at the N terminus (24) and a single hydrophobic membrane-spanning domain toward the C terminus (10). It is clear from the amino acid sequence of gp10 for EHV-4 that neither of these criteria is fulfilled, which would indicate that gp10 is not present as a typical integral membrane protein. Since gp10 shows both immunological and protein sequence homology with the tegument protein VP13/14 of HSV-1, it would seem likely that gp10 is present within the tegument region of EHV-1 and EHV-4.

The presence of a glycosylated protein within the tegument is unexpected, but a protein of this type has been described within the tegument of human cytomegalovirus (3). This protein was found to be glycosylated by radiolabeling with [^3H]galactose and galactosyl transferase, and it was thus considered to be glycosylated in a similar fashion to a number of cytoplasmic and nuclear proteins that have been found in mammalian systems (13). These proteins are glycosylated within the cytoplasm of the cell by specific glycosyl transferases, which add short chains of *N*-acetylglucosamine as O linkages. Examples of such proteins include Sp1, the human RNA polymerase II transcription factor (12), nuclear pore proteins such as p62 (7), and erythrocyte band 4.1 (7, 16). The function of the carbohydrate is not clear, but it may play a role in either nuclear localization or the biological activity of the protein. Since gp10 contains no signal sequence for membrane insertion, it is likely that glycosylation occurs by a similar mechanism, i.e., addition of *O*-glycosyl-linked *N*-acetylglucosamine within the cytoplasm.

The levels of glycosylation of gp10 appear to be higher than those of most characterized nonenvelope glycoproteins. The predicted molecular weight for the B6 gene product is around 97,000; since the mature protein in the virus particle has an estimated molecular weight of 123,000,

this would suggest that carbohydrate residues represent a molecular weight of 26,000. However, no shift in mobility of gp10 was found when the protein was treated with endo- or exoglycanase enzymes (data not shown).

The glycosylation of gp10 raises the possibility that VP13/14 is also glycosylated. If this does occur, the glycosylated protein must be of very low abundance, since when HSV-1-infected cells were labeled in vivo with high levels of [^{14}C]glucosamine and [^3H]glucosamine, no VP13/14 was detected (14, 27). Labeling of HSV-1 with glucosamine and immune precipitation with R220 failed to show any conclusive incorporation of label (data not shown), but it does appear that VP13/14, along with some of the other tegument proteins of HSV-1, binds certain lectins, indicating that carbohydrate is present (30), albeit at a low level.

The immunological cross-reactivity between the proteins VP13/14 and gp10 combined with the homology between the genes UL47 and open reading frame B6 allow identification of UL47 as the gene coding for VP13/14. UL47 (or open reading frame B) has been shown to be an important HSV-1 gene; it decreases the α -*trans*-inducing factor-dependent induction of α genes (20). A second gene, UL46 (or open reading frame A), adjacent to UL47, has been shown to activate the α -*trans*-inducing factor, which in itself is present in the same cluster of genes and is coded for by UL48. It is possible that the function of VP13/14 is shared by gp10, since open reading frame B6 is present in a cluster of open reading frames that share homology with HSV-1 UL46, UL47, UL48, and UL49 (24a). The gene cluster from UL46 to UL49 maps to a position between 0.65 and 0.7 map units on the HSV-1 genome. The EHV U_L region is inverted relative to the HSV U_L region (9), so this gene cluster is a position colinear with the region of EHV-4 that includes open reading frame B6, a position between 0.05 and 0.15 map units.

It is interesting to note that one of the nuclear and cytoplasmic glycoproteins currently identified (Sp1) activates viral and cellular genes. The carbohydrate residues on Sp1 are thought to play a role in the mechanism or regulation of transcriptional control, and all eight RNA polymerase II transcription factors tested from different species were found to be glycosylated to various degrees (12). It is possible, therefore, that glycosylation represents a common feature of proteins that regulate gene expression.

EHV-4 B6
HSV-1 UL47
VZV Gene 11

1	M	D	Q	H	G	V	R	G	G	A	P	I	R	R	P	R	R	S	I	E	T	R	S	N	P	-	-	F	R	A	A	G	N	T	Q	R	-	Y	-		
1	M	Q	S	G	H	Y	R	-	-	P	A	G	R	R	R	R	R	I	S	S	N	Y	T	D	S	P	-	-	R	H	T	H	G	T	R	Y	R	S	T	N	
37	-	T	S	T	P	-	-	R	L	S	Y	R	D	G	L	S	G	R	A	S	S	L	E	P	G	Q	A	H	D	Q	N	E	S	S	T	Q	S	-	-		
36	Y	L	R	A	V	F	R	G	D	D	S	E	T	L	-	-	-	-	-	-	-	-	E	A	L	-	E	E	M	A	G	D	E	P	P	V	R	R	R		
37	W	Y	T	H	P	P	Q	I	L	S	N	S	E	T	L	-	-	V	A	-	V	Q	E	L	L	N	S	E	M	D	Q	D	S	S	S	D	A	S	D	D	
72	-	-	T	S	N	N	Q	P	S	T	S	F	U	G	Y	L	R	R	V	F	S	-	D	D	A	P	A	Q	P	A	P	R	S	R	A	D	F	A	P		
67	E	G	P	R	A	-	-	-	R	R	R	A	S	E	-	A	P	P	T	S	H	R	R	A	S	R	-	Q	R	P	G	P	O	A	A	R	S	Q	S		
74	F	P	G	Y	A	L	H	N	S	T	Y	N	G	S	E	Q	N	T	S	T	S	-	N	H	E	N	R	I	F	K	L	T	E	R	E	A	N	E	E	I	
109	P	P	E	E	D	-	S	S	S	E	E	E	D	E	E	G	P	S	Q	A	P	L	D	E	E	D	Q	L	N	Y	A	D	Q	Y	S	V	G	N	S	S	
102	V	R	G	R	L	D	D	D	D	E	V	P	R	G	P	P	-	-	Q	A	R	Q	G	G	Y	L	G	P	V	D	A	R	A	I	L	G	R	V	G	G	
113	N	I	N	T	D	A	I	D	D	E	G	E	A	E	E	G	-	-	E	A	E	D	A	I	D	D	E	G	E	A	E	E	G	A	E	E	D	A	S	D	
148	D	D	N	E	E	D	Y	L	Q	P	E	V	E	Y	P	T	S	A	E	S	G	E	Y	H	N	S	G	M	F	A	E	E	-	E	-	P	E	S	E	S	
140	S	R	V	A	P	S	P	L	F	L	E	E	L	Q	Y	E	E	D	D	Y	P	E	-	A	-	V	G	-	P	E	D	G	G	G	A	E	E	S	P		
151	I	D	D	E	G	E	A	E	E	G	E	A	E	E	D	A	I	D	D	E	G	E	-	A	E	E	G	-	E	A	E	E	G	E	A	E	E	G	E	A	
186	E	S	D	H	E	N	Y	E	T	Y	E	E	N	D	-	-	T	F	V	I	S	D	D	S	H	R	L	T	R	T	W	L	D	R	S	I	R	L	M	D	
177	V	E	V	L	E	G	N	V	P	G	P	E	L	R	A	A	F	P	L	D	R	L	A	P	Q	V	A	V	W	D	-	E	S	-	-	-	V	R	S	A	
189	E	E	D	A	T	D	D	E	G	E	A	E	E	D	A	A	E	E	D	A	I	D	D	E	G	E	A	E	E	D	Y	F	S	V	S	Q	V	C	S	R	
224	D	A	L	A	O	S	S	E	I	S	K	A	I	T	K	S	T	R	R	L	Y	D	S	Q	F	T	P	G	G	R	G	Y	K	Q	T	E	-	T	P	S	
213	L	A	L	G	H	P	A	G	F	Y	P	C	D	S	A	-	-	G	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
229	D	A	D	E	V	Y	F	T	L	D	P	E	I	S	Y	S	T	-	D	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
263	Q	R	L	V	H	L	S	R	A	G	M	Y	D	S	D	E	I	V	M	T	G	D	Y	-	-	M	E	V	D	D	O	P	N	S	A	Y	Q	S	M	V	
243	K	M	P	A	V	F	F	K	R	-	-	-	-	Q	T	L	Q	Q	G	E	-	A	L	A	W	Y	I	T	G	D	G	I	L	D	L	T	D	R	T	K	
259	D	E	L	N	V	S	K	R	N	C	V	E	P	V	T	L	-	T	G	S	-	M	L	A	-	-	N	N	G	F	G	E	S	W	F	A	M	R	E	C	
301	R	A	I	H	N	P	V	-	A	M	N	P	S	W	E	E	T	I	S	H	M	T	N	T	S	F	S	A	D	I	D	Y	D	I	D	E	L	I	E	M	
278	T	S	P	A	Q	A	M	S	F	L	A	D	A	V	V	R	L	A	I	N	G	U	V	C	G	T	R	L	N	A	E	A	R	G	S	D	L	-	D	D	
295	R	I	R	E	Y	I	T	V	Q	G	L	Y	D	P	I	N	L	R	Y	Q	F	D	T	S	R	M	T	P	P	Q	I	L	R	T	T	P	A	L	P	N	M
340	N	L	A	R	T	P	P	V	F	E	G	L	L	D	S	A	D	F	P	Y	R	L	P	H	L	Y	T	Y	A	T	I	T	Q	D	E	A	Y	E	E	R	
317	R	A	A	-	-	-	-	-	-	E	L	R	R	Q	F	A	S	L	T	A	L	R	P	-	V	G	A	A	V	P	L	L	S	A	G	-	-	-	-	-	
335	T	L	-	-	-	-	-	-	-	G	E	L	L	I	F	P	I	E	F	M	A	Q	P	-	-	I	S	I	E	R	T	L	V	E	D	V	F	L	D	X	
380	G	A	W	S	N	T	Q	A	L	H	G	H	E	Q	S	S	W	P	A	L	V	S	D	Y	S	K	G	G	M	Y	V	S	P	T	Q	E	P	R	G	I	
349	S	P	O	S	G	P	-	D	A	A	V	F	R	-	S	S	L	G	S	L	-	-	L	Y	M	P	G	-	-	-	-	-	-	-	-	-	-	-	-	-	
367	R	A	S	S	K	T	-	N	K	Y	G	P	N	W	M	S	V	Y	A	L	-	-	P	Y	N	A	G	K	M	Y	V	Q	H	I	P	G	F	Y	D	V	
420	-	-	W	R	R	A	L	K	D	A	M	A	L	Q	L	K	L	C	V	L	G	L	T	E	F	V	T	K	R	E	L	T	Q	H	N	S	A	V	F		
381	R	V	A	A	R	Y	A	G	R	M	Y	-	Y	L	A	T	G	A	L	-	L	A	R	F	M	P	D	A	V	R	C	V	T	R	E	A	A	F	F		
404	-	-	S	L	R	A	V	G	Q	G	T	A	-	I	W	H	N	M	I	L	-	S	T	A	A	C	A	I	S	N	R	I	S	H	G	D	G	L	G	F	
458	L	V	D	S	L	R	T	A	K	N	G	Y	L	A	S	R	L	L	V	F	A	W	E	R	R	R	E	T	G	V	R	R	P	A	E	P	L	I	A		
418	L	G	R	V	L	D	V	L	A	V	M	A	E	Q	T	V	Q	W	L	S	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
440	L	L	D	A	A	I	R	I	S	A	N	G	I	F	L	T	R	N	D	N	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
498	L	S	G	V	T	L	L	Q	P	L	P	E	V	S	E	L	E	Q	R	T	-	F	D	I	G	L	R	T	P	Q	S	G	V	F	R	A	F	F			
448	-	V	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
472	L	A	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
537	S	P	L	V	Y	V	A	E	I	R	-	R	A	L	R	D	P	A	I	N	-	-	C	R	Y	V	G	F	H	L	Q	T	S	E	I	Y	L	K	-		
478	S	A	L	G	D	T	A	A	R	R	L	L	A	N	S	G	N	A	V	L	G	A	V	I	A	-	L	N	T	A	L	A	T	I	L	K	-	-	-		
506	G	A	L	A	V	V	P	E	I	R	-	I	A	I	S	E	D	S	T	S	L	-	-	V	R	Y	A	T	G	H	M	E	L	A	E	M	F	L	-	F	
573	A	R	-	-	-	-	-	-	-	A	N	S	A	S	P	G	Y	T	-	K	E	-	E	L	V	A	N	E	A	T	L	-	-	-	-	-	-	-	-		
517	A	R	A	C	G	D	A	N	R	R	R	D	D	A	A	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
542	S	R	-	-	-	-	-	-	-	T	H	S	L	K	P	O	F	T	P	I	E	R	E	N	L	A	S	F	F	T	L	Y	V	T	L	G	G	G	M	-	
605	L	Q	W	I	H	Y	V	A	S	A	-	Q	L	L	S	E	N	D	A	L	K	A	F	R	R	V	S	A	S	I	P	H	A	L	A	P	L	G	S	I	R
552	T	V	V	A	C	R	A	T	A	M	Y	L	A	A	P	Y	H	S	R	S	A	Y	I	A	V	V	C	E	S	L	P	L	R	V	A	C	V	N	L	R	A
574	L	N	M	I	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
644	L	N	D	A	E	F	E	V	L	S	M	P	D	V	M	V	A	R	D	E	T	A	L	S	Q	A	L	F	L	-	G	Y	F	S	V	R	T	A	L	T	
586	Y	G	P	L	T	A	R	T	T	P	A	K	F	-	W	A	D	V	-	R	A	A	E	N	V	D	L	R	-	P	A	S	S	A	P	R	C	A	P	V	
614	L	C	D	L	E	V	L	L	G	E	V	D	-	L	P	T	V	C	E	S	T	A	T	I	A	H	E	L	T	G	Y	E	A	V	R	T	A	A	T		
683	A	C	H	R	D	Y	A	N	E	V	D	G	G	S	K	E	T	V	T	G	L	F	L	G	V	G	L	-	I	I	O	R	L	A	G	H	M	F	-		
623	S	G	T	A	D	P	A	F	L	L	K	D	L	E	P	F	P	A	P	V	S	G	G	S	V	L	G	P	R	V	R	V	V	D	T	H	S	Q	F		
653	N	G	H	I	E	F	A	D	C	Y	K	E	S	E	T	D	L	N	V	S	A	Y	L	G	A	V	L	-	L	L	O	R	V	L	G	H	A	M	L	-	
721	-	-	L	L	M	C	M	A	G	-	A	A	L	T	G	C	S	K	I	A	T	H	S	L	T	P	R	Y	S	L	L	A	D	V							

ACKNOWLEDGMENTS

This work was supported by grants from the Equine Virology Research Foundation and the Wellcome Trust.

REFERENCES

- Allen, G. P., and J. T. Bryans. 1986. Molecular epizootiology, pathogenesis and prophylaxis of equine herpesvirus-1 infections. *Prog. Vet. Microbiol. Immunol.* 2:78-144.
- Allen, G. P., and M. R. Yeargen. 1987. Use of λ gt11 and monoclonal antibodies to map the genes for the six major glycoproteins of equine herpesvirus 1. *J. Virol.* 61:2454-2461.
- Benko, D. M., R. S. Haltiwanger, G. W. Hart, and W. Gibson. 1988. Virion basic phosphoprotein from human cytomegalovirus contains O-linked *N*-acetylglucosamine. *Proc. Natl. Acad. Sci. USA* 85:2573-2577.
- Biggin, M. D., T. J. Gibson, and G. F. Hong. 1983. Buffer gradient gels and 35 S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* 80:3963-3965.
- Cassai, E. N., M. Sarmiento, and P. G. Spear. 1975. Comparison of the virion proteins specified by herpes simplex virus types 1 and 2. *J. Virol.* 16:1327-1331.
- Cullinane, A. A., F. J. Rixon, and A. J. Davison. 1988. Characterization of the genome of equine herpesvirus 1 subtype 2. *J. Gen. Virol.* 69:1575-1590.
- Davis, L. I., and G. Blobel. 1987. Nuclear pore complex contains a family of glycoproteins that includes p62: glycosylation through a previously unidentified cellular pathway. *Proc. Natl. Acad. Sci. USA* 84:7552-7556.
- Davison, A. J., and J. E. Scott. 1986. The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* 67:1759-1816.
- Davison, A. J., and N. M. Wilkie. 1983. Location and orientation of homologous sequences in the genomes of five herpesviruses. *J. Gen. Virol.* 64:1927-1942.
- Engelman, D. M., and T. A. Steitz. 1981. The spontaneous insertion of proteins into and across membranes. The helical hairpin hypothesis. *Cell* 23:411-422.
- Kyte, J., and R. F. Doolittle. 1982. A simple model for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.
- Jackson, S. P., and R. Tjian. 1988. O-glycosylation of eukaryotic transcription factors. Implications for mechanisms of transcriptional regulation. *Cell* 55:125-133.
- Hart, G. W., R. S. Haltiwanger, G. D. Holt, and W. G. Kelly. 1989. Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* 58:841-874.
- Heine, J. W., R. W. Honess, E. Cassai, and B. Roizman. 1974. Proteins specified by herpes simplex virus. XII. The virion polypeptides of type 1 strains. *J. Virol.* 14:640-651.
- Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* 114:193-197.
- Holt, G. D., R. S. Haltiwanger, C. R. Torres, and G. W. Hart. 1987. Erythrocytes contain cytoplasmic glycoproteins: O-linked GlcNAc on band 4.1. *J. Biol. Chem.* 262:14847-14850.
- Honess, R. W., and B. Roizman. 1973. Proteins specified by herpes simplex virus. XI. Identification and relative molar rates of synthesis of structural and nonstructural herpes virus polypeptides in the infected cell. *J. Virol.* 12:1347-1365.
- Lemaster, S., and B. Roizman. 1980. Herpes simplex virus phosphoproteins. II. Characterization of the virion protein kinase and of the polypeptides phosphorylated in the virion. *J. Virol.* 35:798-811.
- McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor. 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. *J. Gen. Virol.* 69:1531-1574.
- McKnight, J. L. C., P. E. Pellett, F. J. Jenkins, and B. Roizman. 1987. Characterization and nucleotide sequence of two herpes simplex virus 1 genes whose products modulate α -trans-inducing factor-dependent activation of α genes. *J. Virol.* 61:992-1001.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Meredith, D. M. Unpublished data.
- Meredith, D. M., J. M. Stocks, G. R. Whittaker, I. W. Halliburton, B. W. Snowden, and R. A. Killington. 1989. Identification of the gB homologues of equine herpesvirus types 1 and 4 as disulphide-linked heterodimers and their characterization using monoclonal antibodies. *J. Gen. Virol.* 70:1161-1172.
- Pereira, L., E. Cassai, R. W. Honess, B. Roizman, M. Terni, and A. Nahmias. 1976. Variability in the structural polypeptides of herpes simplex virus 1 strains: potential application in molecular epidemiology. *Infect. Immun.* 13:211-220.
- Perlman, D., and H. O. Halvorson. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *J. Mol. Biol.* 167:391-409.
- Riggio, M. Unpublished data.
- Riggio, M. P., A. A. Cullinane, and D. E. Onions. 1989. Identification and nucleotide sequence of the glycoprotein gB gene of equine herpesvirus 4. *J. Virol.* 63:1123-1133.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
- Spear, P. G., and B. Roizman. 1972. Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpesvirion. *J. Virol.* 9:143-159.
- Turtinen, L. W., and G. P. Allen. 1982. Identification of the envelope surface glycoproteins of equine herpesvirus type 1. *J. Gen. Virol.* 63:481-485.
- Vieira, J., and J. Messing. 1982. The pUC plasmids, an M13mp 7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19:259-268.
- Whittaker, G. Unpublished data.