JOURNAL OF VIROLOGY, Nov. 1996, p. 8195–8198 Vol. 70, No. 11
0022-538X/96/$04.00+0

The Highly O-Glycosylated Glycoprotein gp2 of Equine Herpesvirus 1 Is Encoded by Gene 71

JANET E. WELLINGTON,[1] GEORGE P. ALLEN,[2] ANDREW A. GOOLEY,[3] DARIA N. LOVE,[4] NICOLLE H. PACKER,[3] JUN X. YAN,[3] AND J. MILLAR WHALLEY[1]*

School of Biological Sciences,[1] and Macquarie University Centre for Analytical Biotechnology, Schools of Biological Science and Chemistry,[3] Macquarie University, Sydney 2109, and Department of Veterinary Pathology, University of Sydney, Sydney 2006,[4] New South Wales, Australia, and Department of Veterinary Science, University of Kentucky, Lexington, Kentucky 40546[2]

Received 3 May 1996/Accepted 22 July 1996

There have been conflicting reports regarding the gene assignment of the high-molecular-mass envelope glycoprotein gp2 (gp300) of equine herpesvirus 1. Here, we provide an unequivocal demonstration that gp2 is encoded by gene 71. gp2 that was purified with a defining monoclonal antibody was cleaved internally to yield a 42-kDa protein encoded by gene 71. Amino acid composition data and N-terminal sequence analysis of a tryptic peptide identified gp2 as the product of equine herpesvirus 1 gene 71 with the SWISS-PROT database. Analysis of gp2's monosaccharide composition and the 42-kDa subunit showed that the high level of O glycosylation occurs on the serine/threonine-rich region upstream of the cleavage site.

Herpesvirus glycoproteins, of which 10 or more are present in both the virus envelope and the plasma membrane of infected cells (15), play key roles in the infectious process and are major targets for host immune responses. Glycoproteins of equine herpesvirus 1 (EHV-1), which is an agent of respiratory disease and abortion, were initially characterized by Turtinen and Allen (19) and were numbered on the basis of relative mobilities on electrophoretic gels. Most of these glycoproteins have homologs in herpes simplex virus type 1 (HSV-1) and have now been assigned to specific genes, a task made easier by the availability of the complete 150-kbp DNA sequence of the EHV-1 genome (18). However, there has been considerable controversy regarding the gene assignment of a major glycoprotein, gp2 (also known as gp300), which has a reported molecular mass in the range of 192 to >400 kDa and contains high levels of O-linked glycosylation together with some N-linked carbohydrate (17, 23). The gene encoding gp2 was originally mapped to the *Eco*RI I fragment in the unique long region of the EHV-1 genome between map units 0.279 and 0.344. This was based on recognition by the gp2-defining monoclonal antibody (MAb) 1G12 of the protein expressed from a phage library of EHV-1 DNA fragments (2). In agreement with this genomic location, Whittaker et al. (22) reported that gp300 was encoded by gene 28, the homolog of HSV-1 UL32. This assignment was based mainly on recognition of a polypeptide expressed from part of EHV-1 gene 28 by the anti-gp300 MAb P19 and on interpretation of sequence data which suggested that the gene 28 product was a novel herpesvirus glycoprotein with multiple membrane-spanning domains.

In contrast, evidence that EHV-1 gp300 is encoded by gene 71 rather than gene 28 was obtained with a gene 71 deletion mutant which was shown to lack the high-molecular-mass protein recognized by MAb P19 and by a specific antiserum prepared against the carboxy terminus of the gene 71 polypeptide (17). The EHV-1 gene 71 is of particular interest in that it is much larger than its positional counterpart gJ in HSV-1, and it appears not to have amino acid sequence similarity to glycoproteins of other herpesviruses (11). It translates after signal cleavage to a predicted 797-residue polypeptide with a molecular mass of 78 kDa containing a serine/threonine-rich N-terminal domain and a C-terminal domain containing all 12 cysteine residues and two potential sites for N-linked glycosylation (18). During EHV-1 infection in cell culture, the gene 71 protein is cleaved internally, dividing the molecule into the two domains (20). A 42-kDa protein recognized by MAb 8B6 was identified by N-terminal amino acid sequencing as a 290-residue C-terminal cleavage product of the gene 71 protein. Here, we describe the use of this MAb in conjunction with the definitive gp2 MAb 1G12 (2), amino acid analysis of purified gp2, and N-terminal sequencing to confirm that this unusual herpesvirus glycoprotein is encoded by EHV-1 gene 71.

Although on Western blots (immunoblots) MAb 8B6 recognized a high-molecular-mass species, similar to the gp2 protein identified by 1G12 (21), we could not be certain that these two MAbs were targeting the same protein. MAb 1G12 did not recognize the 42-kDa 8B6 antigen, possibly because 1G12 was directed against the N-terminal region of the gene 71 product. However, we reasoned that if MAbs 8B6 and 1G12 recognized the same protein, then cleavage of high-molecular-mass gp2 should yield the 42-kDa product, which would then be detectable by MAb 8B6. Since cleavage of the gene 71 protein occurs in RK13 cells during infection with EHV-1, we investigated whether a lysate of uninfected RK13 cells would cleave purified gp2 as follows. EHV-1 (strain HVS25A) gp2 was purified by immunoaffinity chromatography as described previously (20) with MAb 1G12 (2). RK13 cells in phosphate-buffered saline (PBS) were lysed by sonication and added to purified gp2 in PBS containing 0.5% Triton X-100, and the mixture was incubated at 37°C for as long as 12 h. Samples were analyzed by electrophoresis and Western blotting and were compared with purified 8B6 antigen and an extract of EHV-1-infected RK13 cells (Fig. 1). In the presence of the RK13 cell lysate, gp2 was indeed cleaved to produce a subunit which was recognized by MAb 8B6 and which was similar in electrophoretic mobility to the 42-kDa 8B6 antigen, which was previously identified as the C-terminal portion of the gene 71 product (20). Furthermore, high-molecular-mass 8B6 antigen material was recog-

* Corresponding author. Mailing address: School of Biological Sciences, Macquarie University, Sydney 2109, Australia. Phone: 612-850-8200. Fax: 612-850-8245. Electronic mail address: mwhalley@rna.bio.mq.edu.au.

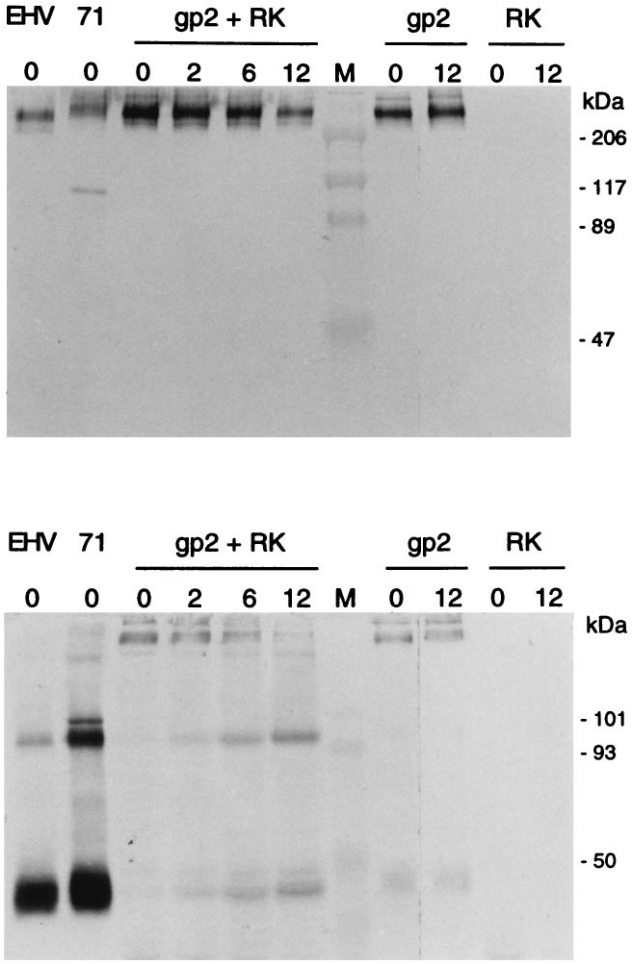

FIG. 1. Western blots from SDS-PAGE of EHV-1 gp2 affinity purified with MAb 1G12. EHV, RK13 cells infected with EHV-1; 71, gene 71 protein material purified with MAb 8B6; gp2 + RK, purified gp2 incubated with extract of RK13 cells; gp2, gp2 incubated without added RK13 cells; RK, RK13 cells without added gp2; M, molecular mass markers. The numbers above the lanes are incubation times at 37°C (in hours). (Upper panel) Detection with MAb 1G12 (anti-gp2) (2); (lower panel) detection with MAb 8B6 (anti-gene 71) (20).

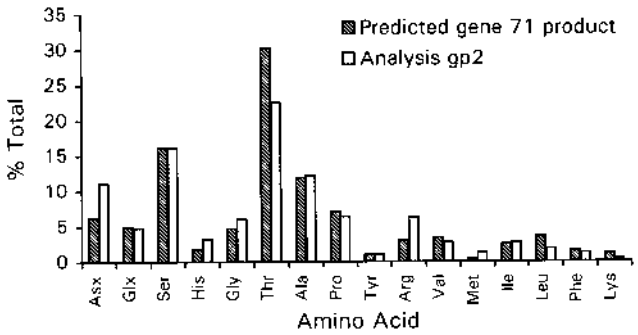

FIG. 2. Comparison of the experimentally determined amino acid composition of purified gp2 with the amino acid composition of the predicted translation product of EHV-1 gene 71 (18) after removal of the N-terminal signal cleavage peptide. Asx, aspartic acid plus asparagine; Glx, glutamic acid plus glutamine. Cysteine and tryptophan were not included in the analysis.

nized by 1G12 (anti-gp2) and gp2 was recognized by MAb 8B6. A third MAb, 2E3, also recognized both gp2 and the high-molecular-mass 8B6 antigen (results not shown). These results provide compelling evidence that MAbs 1G12 (anti-gp2) and 8B6 (anti-gene 71 product) are directed against the same protein, and, therefore, that gp2 must be encoded by EHV-1 gene 71.

Additional evidence for the assignment of gp2 to gene 71 was obtained from direct amino acid analysis. gp2, which had been immunoaffinity purified with MAb 1G12, was further purified by gel electrophoresis and passive elution of the high-molecular-mass band overnight at 37°C into 100 mM acetate buffer (pH 8.5) containing 0.1% sodium dodecyl sulfate (SDS). The gp2 in the eluate was then centrifuged onto a polyvinylidene difluoride membrane, subjected to vapor-phase hydrolysis, and analyzed with the automated AMINOMATE system (GBC Scientific Equipment, Victoria, Australia) (13, 24). The amino acid composition thus determined for the purified gp2 was notable for its high serine and threonine content and for an overall amino acid profile which closely matched that of the predicted gene 71 polypeptide after signal cleavage (Fig. 2). The experimental amino acid composition of gp2 was matched against all 78 EHV-1 SWISS-PROT entries (18), with the SWISS two-dimensional polyacrylamide gel electrophoresis (2D PAGE) computer program on the World Wide Web at http://expasy.hcuge.ch//. The first-ranked entry was the open reading frame of gene 71 of EHV-1 (accession no. P28968), with a score of 103, while the second-ranked EHV-1 protein gave a score of 311. Gene 28 (accession no. P28952) was ranked 23rd. The gene 71 product was also ranked first in the entire SWISS-PROT database. According to the SWISS 2D PAGE score and rank pattern, gp2 is the likely product of gene 71.

Attempts to sequence the N-terminal amino acids of purified high-molecular-mass gp2 were unsuccessful, probably because of a blocked N terminus. However, despite the large mass of gp2, affinity-purified protein was reduced and alkylated with acrylamide and subjected to SDS-PAGE and in situ digestion with trypsin according to the method described by Moritz et al. (10). The resultant peptides were separated by reverse-phase chromatography (Sephacil C8 [2 by 100 mm]) on a SMART high-performance liquid chromatography column, and one peptide (retention time, 33 min) was subjected to Edman degradation as previously described (20). The sequence was determined as XXXIPCPGQQ, which, when run through the database, again matched the gene 71 translation product (18) and corresponded to the peptide STTIPCPG QQR from residues 656 to 666, which was flanked by a lysine residue at 655 and was terminated by arginine, in agreement with a product of trypsin digestion. This internal amino acid sequence provides final, conclusive evidence that gp2 is the product of gene 71. These results for gp2 are consistent with the conclusions of Sun et al. (17) for gp300 and, therefore, also confirm that gp2 and gp300 are one and the same protein (23). It is possible that the earlier links to the *Eco*RI I fragment and gene 28 location (2, 22) were due to a cross-reacting epitope(s) recognized under the conditions and with the reagents used, which in each case involved the EHV-1 polypeptide expressed as a fusion product in *Escherichia coli*. Comparisons of amino acid sequence and predicted antigenic regions of the open reading frames of EHV-1 genes 71 and 28 and other genes in the *Eco*RI I fragment do not reveal any obvious potentially cross-reacting epitopes. Therefore, in the absence of further information on the epitope specificities of MAbs P19 and 1G12, the nature of any cross-reaction remains speculative.

To further characterize the gene 71 products, proteins pu-

TABLE 1. Monosaccharide compositions of EHV-1 gene 71 products

Monosaccharide	Amt (pmol)[a]	
	42-kDa protein	Uncleaved protein
N-acetylglucosamine	216 (29)	120 (7)
Mannose	358 (48)	140 (8)
N-acetylgalactosamine	28 (4)	318 (19)
Galactose	140 (19)	384 (23)
Fucose	ND (ND)	275 (16)
Sialic acids		
N-acetylneuraminic acid	ND (ND)	266 (16)
N-glycolylneuraminic acid	ND (ND)	150 (9)

[a] Values in parentheses are percentages of total. The amount of each glycoprotein used for this analysis was not known; the yield of each monosaccharide cannot be compared between glycoproteins (except as a percentage of the total monosaccharides detected). Values for glucose are not given because the amido black stain is contaminated with glucose. ND, not detected.

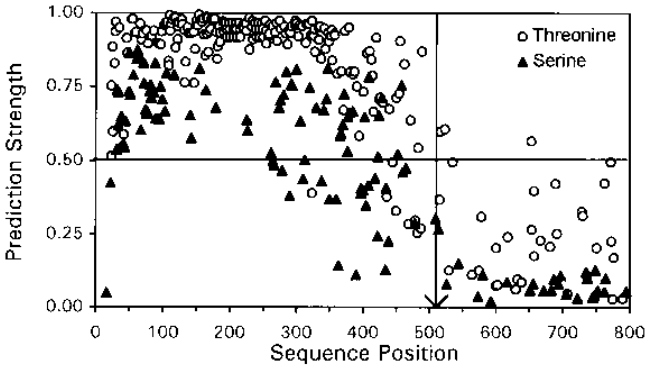

FIG. 3. O-glycosylation site prediction for the EHV-1 gene 71 protein (8). All serine and threonine residues with a prediction strength of more than 0.5 are likely glycosylation sites. Vertical arrow, internal cleavage site in the gene 71 protein (20).

rified with MAb 8B6 were analyzed for monosaccharide composition. Purified 8B6 antigen was electrophoresed and transferred to a polyvinylidene difluoride membrane, and both the 42-kDa and the high-molecular-mass proteins were subjected to a two-step hydrolysis procedure which sequentially removed sialic acids and neutral and amino sugars (14). The hydrolysates were dried under a vacuum and were analyzed by high-pressure anion exchange chromatography (Dionex Carbopac PA1) and electrochemical detection. Quantitation was by the addition of internal standards (2-deoxyglucose for neutral and amino sugars and lactobionic acid for sialic acids). Assuming the usual characteristic monosaccharide composition of glycoproteins, the results in Table 1 show that the uncleaved gene 71 product was modified predominantly with O-linked oligosaccharides (*N*-acetylgalactosamine, 19%) and with some complex N-linked oligosaccharides (mannose, 8%). The low proportion of *N*-acetylglucosamine (7%) relative to the high proportions of the monosaccharides galactose, fucose, and sialic acids indicated an abundance of O-linked oligosaccharides. This direct chemical analysis is consistent with previous studies of gp300 with glycosidases, lectin binding, and inhibitors (17, 23), which indicated high levels of O glycosylation and a small amount of N glycosylation. The 42-kDa cleavage product was modified predominantly with complex N-linked oligosaccharides, as indicated by the high proportion of *N*-acetylglucosamine and mannose, with little indication of O-linked oligosaccharides (*N*-acetyl galactosamine, 4%) (Table 1), indicating that the high level of O glycosylation on the uncleaved gene 71 product occurs on the serine/threonine-rich region upstream of the cleavage site. These experimental findings are in agreement with a predictive O-glycosylation analysis of the gene 71 sequence obtained through the Internet address NetOglyc@cbs.dtu.dk, by using the algorithm described by Hansen et al. (8), which indicates that a high proportion of the threonine and serine residues N terminal to the cleavage site are likely to be O glycosylated (Fig. 3).

The function of this apparently unique herpesvirus glycoprotein has not yet been fully elucidated; however, the data indicate that gp2 is a mucin-like glycoprotein with abundant O-glycosylation sites in the N-terminal part of the molecule (8). Most cell surface mucins have a small number of N-linked chains, and some are proteolytically cleaved to form one serine/threonine-rich subunit and one N-glycosylated subunit (5). Such cleavage into two domains also occurs in gp2 (21), although full-length gp2 is also present in virions (19). It has been suggested that a heavily O-glycosylated gene 71 protein may form extended rods from the virion surface and could have considerable effects on properties of the virion, including antibody accessibility and cellular binding (9). Although it has been shown to be dispensable for growth in cell culture (16), a gene 71 deletion mutant was less virulent than wild-type EHV-1 (16). In addition, a number of monoclonal antibodies to the gene 71 products do appear to reduce plaque size and to affect cell-cell fusion (7, 22).

In an attempt to obtain some additional information on the possible origin of this gene and in light of the confirmation of gp2 as a mucin-like product with two domains, we searched the SWISS-PROT database using the C-terminal fragment of gene 71 but did not identify any protein with large homologous sequence. However, a short string of 44 amino acid residues (503 to 547) immediately downstream of the gp2 cleavage site shared 36% identity with residues 775 to 819 of the 907-amino-acid Epstein-Barr virus membrane antigen gp350 (4) (accession no. P03200). Although there was no convincing similarity over the rest of the sequence, gp350 is also heavily glycosylated (in this case, by N-linked sugars), with an amino acid composition and a pI similar to those of EHV-1 gp2. The homologous amino acid string was not close to the B-lymphocyte CR2 receptor binding sequence in gp350 (12); therefore, no specific functional significance could be imputed for EHV-1 gp2 from these results. A search which included the N-terminal region, with its high serine and threonine content, showed homology with many proteins, some of which are mucins. The full gene 71 sequence is not yet available for the genetically closely related EHV-4. However, a partial N-terminal sequence of 116 amino acids reported for the EHV-4 gene 71 (11) had an unexpectedly low (45%) amino acid homology with the same region of the EHV-1 protein, possibly reflecting a less-stringent requirement for exact sequence in highly O-glycosylated domains compared with other functional regions. Different relative mobilities of gp2 observed for EHV-1 and EHV-4 (1, 25) are presumably due to such sequence differences. In the Kentucky A strain of EHV-1, inspection of the sequence of the unique short region (6) reveals that this gene 71 has a large deletion, which would result in the removal of most, but not all, of the serine/threonine-rich domain while the C-terminal domain would be retained. Experiments aimed at verifying this postulate are in progress. The suggestion that the Kentucky D strain may also have a shortened gene 71 product of only 198 amino acids (3) is probably incorrect since, when possible frameshifts in the published upstream sequence are taken into consideration, the partial amino acid sequence is identical to that re-

ported for gene 71 in EHV-1 strain Ab4 (18). In any case, we would not expect a short form of gene 71 to occur in the Kentucky D virus, since this strain was used for the original determination of gp2 as a high-molecular-mass protein (19). We now believe that this would require an essentially full-length gene 71, including sequence encoding the serine/threonine-rich N-terminal domain. The conclusive assignment of gp2 (gp300) to gene 71, therefore, resolves a number of apparent anomalies and makes it possible to investigate with confidence the functional and structural properties of this unusual herpesvirus glycoprotein.

J.M.W. and D.N.L. acknowledge support from the Australian Research Council (grant no. A09600826) and the Rural Industries Research and Development Corporation (grant no. UMA14A).

Thanks are due to M. Molloy and M. Lawson for assistance with amino-terminal sequencing and sugar analysis; to CSL Ltd. for supply of media and serum for cell culture; to G. Lawrence for helpful discussions; and to E. Smith for reading the manuscript.

REFERENCES

1. **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.
2. **Allen, G. P., and M. R. Yeargan.** 1987. Use of λgt11 and monoclonal antibodies to map the genes for the six major glycoproteins of equine herpesvirus 1. J. Virol. **61:**2454–2461.
3. **Audonnet, J.-C., J. Winslow, G. Allen, and E. Paoletti.** 1990. Equine herpesvirus type 1 unique short fragment encodes glycoproteins with homology to herpes simplex virus type 1 gD, gI and gE. J. Gen. Virol. **71:**2969–2978.
4. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, P. S. Tuffnell, and B. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310:**207–211.
5. **Carraway, K. L., and S. R. Hull.** 1991. Cell surface mucin-type glycoproteins and mucin-like domains. Glycobiology **1:**131–138.
6. **Colle, C. F., C. C. Flowers, and D. J. O'Callaghan.** 1992. Open reading frames encoding a protein kinase, homolog of glycoprotein gX of pseudorabies virus, and a novel glycoprotein map within the unique short segment of equine herpesvirus type 1. Virology **188:**545–557.
7. **Edington, N., C. G. Bridges, and S. C. Broad.** 1987. Rapid diagnosis and characterization of equid herpesvirus 1 using monoclonal antibodies. Tieraerztl. Prax. **2 (Suppl.):**37–40.
8. **Hansen, J. E., O. Lund, J. Engelbrecht, H. Bohr, J. O. Nielsen, J. E. S. Hansen, and S. Brunak.** 1995. Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-galNac-polypeptide N-acetylgalactosaminyltransferase. Biochem. J. **308:**801–813.
9. **McGeoch, D. J., B. C. Barnett, and C. A. Maclean.** 1993. Emerging functions of alphaherpesvirus genes. Semin. Virol. **4:**125–134.
10. **Moritz, R. L., J. Eddes, H. Ji, G. E. Reid, and R. J. Simpson.** 1994. High-speed chromatographic separation of proteins and peptides-application to rapid peptide mapping of in-gel digested proteins. J. Protein Chem. **13:**486–487.
11. **Nagesha, H. S., B. S. Crabb, and M. J. Studdert.** 1993. Analysis of the nucleotide sequence of five genes at the left end of the unique short region of the equine herpesvirus 4 genome. Arch. Virol. **128:**143–154.
12. **Nemerow, G. R., R. A. Houghton, M. D. Moore, and N. R. Cooper.** 1989. Identification of an epitope in the major envelope protein of Epstein-Barr virus that mediates viral binding to the B lymphocyte EBV receptor (CR2). Cell **56:**369–377.
13. **Ou, K. L., M. R. Wilkins, J. X. Yan, A. A. Gooley, Y. Fung, D. Sheumack, and K. L. Williams.** 1996. Improved high-performance liquid chromatography of amino acids derivatised with 9-fluorenylmethyl chloroformate. J. Chromatogr. **723:**219–225.
14. **Packer, N. H., M. R. Wilkins, O. Golaz, M. A. Lawson, A. A. Gooley, D. F. Hochstrasser, J. W. Redmond, and K. L. Williams.** 1996. Characterization of human plasma glycoproteins separated by two-dimensional gel electrophoresis. Bio/Technology **14:**66–70.
15. **Spear, P. G.** 1993. Entry of alphaherpesviruses into cells. Semin. Virol. **4:**167–180.
16. **Sun, Y., and S. M. Brown.** 1994. The open reading frames 1, 2, 71, and 75 are nonessential for the replication of equine herpesvirus type 1 *in vitro*. Virology **199:**448–452.
17. **Sun, Y., A. R. Maclean, D. Dargan, and S. M. Brown.** 1994. Identification and characterization of the protein product of gene 71 in equine herpesvirus 1. J. Gen. Virol. **75:**3117–3126.
18. **Telford, E. A. R., M. S. Watson, K. McBride, and A. J. Davison.** 1992. The DNA sequence of equine herpesvirus 1. Virology **189:**304–316.
19. **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.
20. **Wellington, J. E., A. A. Gooley, D. N. Love, and J. M. Whalley.** 1996. N-terminal sequence analysis of equine herpesvirus 1 glycoproteins D and B and evidence for internal cleavage of the gene 71 product. J. Gen. Virol. **77:**75–82.
21. **Wellington, J. E., D. N. Love, and J. M. Whalley.** 1996. Evidence for involvement of equine herpesvirus 1 glycoprotein B in cell-cell fusion. Arch. Virol. **141:**167–175.
22. **Whittaker, G. R., W. A. Bonass, D. M. Elton, I. W. Halliburton, R. A. Killington, and D. M. Meredith.** 1992. Glycoprotein-300 is encoded by gene-28 of equine herpesvirus type 1—a new family of herpesvirus membrane proteins. J. Gen. Virol. **73:**2933–2940.
23. **Whittaker, G. R., L. A. Wheldon, L. E. Giles, J. M. Stocks, I. W. Halliburton, R. A. Killington, and D. M. Meredith.** 1990. Characterization of the high M_r glycoprotein (gp300) of equine herpesvirus type 1 as a novel glycoprotein with extensive O-linked carbohydrate. J. Gen. Virol. **71:**2407–2416.
24. **Wilkins, M. R., C. Pasquali, R. D. Appel, K. Ou, O. Golaz, J. C. Sanchez, J. X. Yan, A. A. Gooley, G. Hughes, I. Humphery-Smith, K. L. Williams, and D. F. Hochstrasser.** 1996. From proteins to proteomes—large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Bio/Technology **14:**61–65.
25. **Zheng, M., D. N. Love, and M. Sabine.** 1995. High molecular weight polypeptide bands specific for equine herpesvirus 4. Vet. Microbiol. **46:**203–211.