Comparison of the Bovine Herpesvirus 1 gI Gene and the Herpes Simplex Virus Type 1 gB Gene

J. CHARLES WHITBECK,* LEONARD J. BELLO, AND WILLIAM C. LAWRENCE

Laboratory of Microbiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6049

Received 11 February 1988/Accepted 28 May 1988

In a previous report, we localized the gene for a 130-kilodalton envelope glycoprotein (gI) of bovine herpesvirus 1 (BHV-1) to a 3.6-kilobase *HpaI-KpnI* restriction endonuclease fragment from the long unique region of the BHV-1 genome (map position 0.405 to 0.432) and showed that a herpes simplex virus 1 (HSV-1) glycoprotein B (gB) probe uniquely hybridized to this BHV-1 restriction fragment. Here we present the complete nucleotide sequence of the BHV-1 gI gene and the predicted 932-amino-acid sequence of the gI primary translation product. Comparison with the published nucleotide sequence of the HSV-1 (KOS) gB gene (D. J. Bzik, B. A. Fox, N. A. DeLuca, and S. Person, *Virology* 133:301-314, 1984) reveals a similarity of 56.3% at the nucleotide level and 45.9% at the amino acid level. Upstream of the proposed gI coding region are potential mRNA transcriptional promoter elements including a TATA box and multiple Sp1 binding sites (GC boxes). Downstream of the gI coding region are two sequence elements associated with mRNA cleavage and polyadenylation (AATAAA and a GT-rich region roughly 30 nucleotides further downstream). Like HSV-1 gB, the predicted gI amino acid sequence exhibits two broad hydrophobic regions likely to represent a transient amino-terminal signal sequence and a transmembrane anchor domain (near the carboxyl terminus). Additional features shared with gB include 6 potential N-linked glycosylation sites and 10 highly conserved cysteine residues in the gI extracellular domain. Two regions of nonsimilarity between gI and gB are a centrally located 22-amino-acid region of gI for which there is essentially no gB counterpart and the transient amino-terminal leaders which differ in both size and sequence. The hydrophobic signal sequence of the gI leader, unlike that of gB, is preceded by an unusually large region of predominantly hydrophilic amino acids. The unusual length of the gI leader may result from an overlap between that portion of the gI coding region and a potential upstream coding region.

Bovine herpesvirus 1 (BHV-1) is an alphaherpesvirus (24) which is associated with a number of clinical manifestations in cattle. BHV-1 infection can result in severe respiratory tract infections, vulvovaginitis, abortions, conjunctivitis, encephalitis, and generalized systemic infections (8).

The BHV-1 genome is a linear double-stranded DNA molecule of 137 kilobases (kb). The genome can be divided into a long and a short unique region, with the latter being bounded by inverted repeat sequences. Thus, the two observed isomeric forms of the BHV-1 genome result from inversion of the short, relative to the long, unique region.

The BHV-1 virion is surrounded by a lipid envelope which bears a number of virus-encoded glycoprotein species. These glycoproteins are thought to mediate several key steps during the early stages of infection, including viral attachment and penetration of the host cell. At later times in infection, newly synthesized viral glycoproteins are also present on the surface of infected cells (28).

One of the best-characterized BHV-1 glycoproteins is the gI molecule (nomenclature proposed in reference 30; this glycoprotein has also been referred to as GVP 6/11a/16 [29, 31] and gI30 [14]) which exists in the virion as a 130-kilodalton (kDa) form. Analysis of this glycoprotein by sodium dodecyl sulfate-polyacrylamide gel electrophoresis under reducing conditions reveals that is actually a complex of three distinct glycoproteins of 130, 74, and 55 kDa (14, 17, 31). Pulse-chase studies have shown that gI is initially synthesized as a single glycosylated form which undergoes several modifications, including cleavage of the polypeptide

chain (although some molecules appear to remain uncleaved) (30). Thus, the 74- and 55-kDa forms seen under reducing conditions represent disulfide-linked components of the cleaved form of gI.

Recently, this laboratory reported the genomic map position of the BHV-1 gI gene on the basis of marker rescue experiments (14). The site of the mutation in a BHV-1 mutant coding for an antigenically altered gI was localized to within a 3.6-kb *HpaI-KpnI* restriction fragment located at map units 0.405 to 0.432. Evidence for DNA sequence similarity between BHV-1 gI and herpes simplex virus type 1 (HSV-1) gB was provided by Southern blot experiments where BHV-1 restriction fragments were probed with nick-translated HSV-1 gB DNA. The gB probe uniquely hybridized to the same 3.6-kb BHV-1 fragment identified by marker rescue as containing the mutation responsible for the antigenically altered form of gI.

In this report, we extend our previous work by presenting the complete nucleotide sequence and deduced translation of the BHV-1 gI gene. Our results show that the similarity with HSV-1 gB is 56.3% at the nucleotide level and 45.9% at the amino acid level. We also show that a potentially functional upstream coding region overlaps the gI coding region by 47 codons and may affect the character of the gI amino-terminal leader as a result. Evidence for amino acid sequence similarity between BHV-1 gI and the apparent homologs of other herpesviruses is presented in the form of matrix analyses.

MATERIALS AND METHODS

Virus and cell culture. The Colorado-1 (Cooper 1) strain of BHV-1 was obtained from the American Type Culture

* Corresponding author.

HSV-1	GTCAACGGGCCCCCTTTTGATCACTCCACCACAGCTTCGCCAGCCCCCAACA	-493	GCTGAAACGGGCAACGCCGACCCGACGAGC---CGGGGCTGGCACACCACGACCTCAAG	759
BHV-1	GTTAACGGGCGGCTCTTTAGCGCAGCGGCACAACCTTCGCGAGCGCAAACA	-329	GCTGAA---GCCTGCGGGCTGAGCGCGCCGGGTGCGGGGCTGGCACACGACGACGATGTG	792
	Hpa I			
	CCGCGCTGATTACAGCTCGAGAAGTGGGGCTCTGCCGACCTGAAGGAGGAGCTCGCCGG	-428	TACAACCCCTCGGGGTGAGGGCTTCC---ACCGGTACGGGACGACGTAACCTGCATCGTCGA	821
	CCGCGCTGTACTTTAGCTGGAAGAACTGGGGCTGCTCCGCACTCAAGGAGGAGCTGGCGCG	-264	TACACGG---CGCTGGGCTCGGGCGGGCTCTACCGACGGGACCTCTGTGAACCTGCATCGTGA	854
	GC CAAT			
	TTCATCATGGGGCGG-GGGGCTCGGGTCTGATTGGGCGCTCAGCAATTTACAGAGTTTTACT	-364	GGAGGTGACGCGCGCTCGGTGTACCCGTACGACGAGTTTGTGCTGGGACTGGGACTTTGTGT	886
	TTTATGCTGGCGCGCGGGGGCGGGTGG-----GGCGTAAGCGACTTCAGACAGTTTT---	-209	AGAAAGTGAGGCGCGCTCGGTGTACCCGTACGACTGTTGCGGCTCTGACCGGGGACATTATCT	919
	GC GC			
	GTTTTGACGGCATTTCCGGAATAACGCC-----CACT-----CAGCGCGCGCTGGCGATAT	-311	ACATGTCCCCGTTTTACGGTACCAGGAGGGGTGCGACACCGAACACACACG-TACGCCCGGA	950
	-TTTGCTTCGCAT---CCGCGCGGGCGGGGCTCAGCGCGCGCGCTCGCTGGCAATAT	-147	ACATGTGCCCTTTTACGGGCTGCGGAGGGCGCGCACCGCGAGCACACAGGCTACTCGCCGA	984
	GC TATA			
	ATTCGCGAGCTGATTATCGCAACCACTCTTTGCTCGGTCTACCGGTGCGGGAGCTCGAGTT	-246	CCGCTTCAAGCAGGTGACGCGCTTCTACGCGCGACCTACCAACGAGCCCGGGCCACGGC--	1013
	ATTCGCGAGCTCGTTCTGCGCGCGCGCTCTTTGCTGCGCTCTTCCACTGCGGACGCGTCCGCT	-82	GCCTTTCAGCAGATCGAGGGCTACTACAAGCGGACATGACCAG---GGCGGCGCTCAAGGA	1047
	GCGCGCGCGGACTCGACG-CGCGCGACCTCGAAGGCTTACCGTTACCGCGCGGCTATAT	-182	GCCGACACCCGGAACCTGCTCACGACCCCAAGTTACCGTGGCTGGGACTGGTGCCAAAGC	1078
	GCTGCTGCGGACGAAACGGCGCGGGCGGAGCGGCGGAGTGTGTCCAG-CGGCGCTAC	-18	GCCGCTCTCGCGGAATTTTTCGTACACAGCACGTGACGGTAGCTGGGACTGGGTGCCAAGC	1112
	CTCACGTACGACTCCGACTGTCCGCTGGTGGCCATCGTCGAGA-----GCGCCCCGACGGC	-125	GCCGCTGGTCTGCACCATGACCAAGTGGCAGGAAGTGGACGAGATGCTGCGCTCCGAGTACGGC	1143
	CTGACCTACGAGGATCATGGCCGCTCGCGCGGTGTGAACGCGCGCGGGCGCGGAGACGG-	47	GCAAAACGCTGTGCTGCTGCGCAAGTGGCGCGAGGCGGACGAAATGCTGCGAGACGAGACGGC	1177
	TGTATCGGCCCGCGTGGTGTGTCTACGACCGGACGCTTTTCTGATCCTCTACTCGGTCT	-60	GGTCTCTTCGATTCTCTCTCGACGCCATATCCACCACCTTACCACCAACCTGACCGAGT---	1204
	----TCGGCAGGACAGCTGTCTATCTACGACCGGAGCTGTCTCGCTGCTCTACGCGGTCT	108	GGGAACCTTCGCTTACGCGCCGCTCGCTCTCGCGACCTTTG-----TGAGCGACAGCCA	1233
	CCAGCACCTCGCCCCAGGCTACCTGACGGGGGACGACGAGGGCCCCGAGTCCCGCATGAC	6	-ACCGCTCTCGCGC-----GTGGACTGGGGGACTGCATCGGAAGGACGCCCGACGCGC	1260
	GACGCGCT-----	117	CACCT---TCGCTTGCAAGATGTGCGCTGAGCGACTGCTGATGAGAGGCGGAGGCGCG	1294
	CAGGGCGCCCCCTCGTGGGGCGCGGTGTTCTGCTGATGGCGCTCTTGGGGTTGACGCTGG	71	ATGGACCGCATCTTCGCCCGCAGGTACAACGCGACGACATCAAGGTGGGCCAGCCGAGTACTA	1325
	-----		GTGAGCGCGCTTACCGGAGCGCTACAACGCGACGACGCTGCTGCGGCGAGCTTGAGAGCTA	1359
	GGTCTGTGGTGGCTCGCGGCTCCGAGTTCGCCGGCACGCTGGGT---CGCGCGCGGAC	132	CCTGGCCAAATGGGGGCTTTGTGATCGGTACGACCCCTTCTCAGAACACGCTCGCGGAGCTGT	1390
	-----GGCGCGCGCG-----GGCGCGCGCGCTAGCGCTGCGCTGCTA	159	CCTGGCGCGCGCGGCTTTGTGTTGGCTTCCGGC-GATGCTCAGAACAGCTGGCCAAGCTGT	1423
	CAGGCGGCAACGGGGGCTTGCCTCGGCGCGCGCGCGCTT-----GGCGC---CGCC	186	ACGTGCGGGAA-----CACCTCGAGAGCAGAGCGCAA-----	1424
	TGGGCG---ACGTGGGCTTGTGCT---GGCGCGCGCGCGCGGGGACGCGGACGACGCG	218	ACCTGACGAGCTGGCGCTCGAAGCGCACGCTCGAGGGGCTGTTGCGCGCGCGCGCGCCAAAG	1488
	CCAACGGGGGACCCGAACCGAAGAAAGCAAAACCGAAAAACCAACGCCACCAC-----	244	----GCCCCAAACCCACGC-----CCCGCGCGCGGG-----	1455
	CC--CGCGCGCGCGCG--CGAAGAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	279	CCGGGCGCGGGCGCGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGGGCGCGGCAACGG	1553
	-GCCCGCGCGCGGACGACGCGACCTGCGCGCGGGGACGCGACCTGCGGAGCAGCTGCGGGA	308	----GCCAGCGCAACGCTCGGTGGAGCGCATCAAGACCCTCTCTCATGAGTTGCGCGGCG	1516
	GGCCCCGACGCGGACGCGCGCGCGCGCGCGCGGACGACGAGCTGCGCGCGCGCTCGGCT	344	GCCCGCGCGCGCGCGCGCGCGCGCGCGGGGCTGACTACCGTAGCTGCGCGAGTTGCGCGCG	1618
	CATCAAGGCGGAGAACACCGATGCAAACTTTTACGTGTGCCACCCCGACGGGCGCGCGTGG	373	TGCAGTTTACGTACAACACATACAGCGCATGTCAACGATATGTTGGGCGCGTTGCCATCGCG	1581
	CGCGCAGGCGCGGGGAAACTCGCGCTTCTTGTGTGCCCGCGCTCGGGCGCGCGGTTGG	409	TGCAGTTACCTACGACACATCCAGGACACGTAACACCATGTTTCAAGCGCTGGCCACGCTC	1683
	TGCAGTTGAGAGCGCGCGCTGCGCGCGCGCGCGCGGAGGTCAGAACTACAGGAGGGCATC	438	TGGTGCAGCTACAGAAACAGAGTACGCTGTGGAACGAGGCGCGCAAGCTGAACCCCAACGC	1646
	TCCGGCTCGCGCGCGCGCGCGCTGCGCTGAGTACGGGCTCGGGCGGAACTACAGGAGGGCATC	474	TGGTGCCTGCTGAGAAACAGGAGCGCGCTGTGGGCGAGGCGGCTAAGCTCAACCCACGCG	1748
	GCGTGTCTTCAAGGAGAACCTGCCCCGTACAAGTTCAAGGCCACCATGTACTACAAAGAGCT	503	CATCGCTCGGTACCGTGGGCGGGGAGGCGCGGATGCTGCGGACGTTGATGGCGCTCT	1711
	GGCGTCATTTACAAGGAGAACCTGCGCGCTACAGTTCAAGGCTACATTTAC---AAAAAGCT	536	GGCGGCGAGCGCTGCGTGGACCGCGCGCGCGCGCGCATGTTGGGGACGCCATGGCGTGA	1813
	CACCGTTTGCAGGTGTGGTTGCGGCCACGCTACTCCGATTTATGGGATCTTTGAGGACCGG	568	CCAGTGCCTGCGGCTCGCGCGCGGACACGATGCTCCAAACCTGATGCGCATAGCTCGCGG	1776
	GATCGTGACACGACTGGGCGGCGAGCAGTACGCGGCTTTACAAACAGTACAGGACCGCG	601	CGTACTGCCACGAGCTGGGCGAGGGGCGGCTGTTATCAGAACTCGATGCGCG-----CGCCC	1872
	CCCCGTTCCCTTTCAGGAGGTGATCGACAAGATCAACGCGCAAGGGGTCTGTGCTGCTACGGCC	633	CCGGGGCGCTGTACAGCGCGCCCTTGGTCACTTTCGCTGAGAACGAGGCGCGTTGGTCTGA	1841
	TGCCCTGGGATGGGCGAGTACGAGACTGTTGGACAAGAGTGGCGCTGCTTTCAAGAGCC	666	GGCGCGTGTGCTACAGCGCGCGCGCGCTCTCTTTCCTTGGCAACGAGAGGAGCGGCTGA	1937
	AAGTACGTGCGCAACACTGGAGACCCGCGTTTACCAGGACGACGACGAGACGACATGGA	698	GGGGCAGCTGGGGAGAACAGAGTGCAGCTGACGCGCGATGCGATGAGCGGTGACCGTGG	1906
	GAGTACGTGCGAGCGGGCGAAGGTGGTGGCTTTGACCGGACGACGACCCCTGGGAGGCGCG	731	GGGCCAGCTCGCGAGGACAAAGAGTGTGCGGGCGCGAGCTGTTGGAGCCCTGACCGCCA	2002

GACACCGCGCTACTTCACCTTCGGTGGGGCTACGTGACTTCGAGGAGTACGCTACTCCAC	1971
ACCACAAGCGCTACTTCGCTTTGGCGCGGACTACGTGACTACGAGAACTACGCTACGTGCGG	2067
CAGCTGAGCGCGCGACATCACCACCGTCAGCACCTTCATCGACCTCAACATCACCATGCTGGA	2036
CGGGTCCGCTCGCGAGCTGGAGGTGATCAGCACCTTTGTGGACCTAAACCTCAGGTTCTGGA	2132
GGATCAGAGTTTGTCCCTCGAGGTGTACACCGCCACGAGATCAAGGACAGCGCTGCTGG	2101
GGACCGGAGTTCTTGGCGCTAGAAAGTGTACACGCGCGGAGCTCGCCGACACGGGTCTGCTCG	2197
ACTACACGAGGTCAGCGCCGCAACGAGTGCACGACCTGCGCTTCGCGACATCGACACGGTC	2166
ACTACACGAGATACAGCGCGCAACGAGTGCACGAGCTCGGTTCTACGACATTGACCGCGTG	2262
ATCCACGCCGACGCCAACGCCCATGTTGCGGGCTGGGCGGCTTCTCGAGGGGATGGGCGA	2231
GTCAGAGCGGCAATATGGCCATCATGCGAGGGCTCGCCAACCTCTTTTCAGGGCTGGGCGC	2327
CCTGGGCGCGCGTGGCAAGGTGGTATGGGCTGCGGGCGGCTGGTATCGCGCTGCTCGG	2296
CGTGGGCGGCGTGGGCGCGGTGGTGGGCGCGGCTGCGCGCTCTCGACCGTGTGCGG	2392
GCGTGTCTCTTCTATGTCACACCTTTGGGGCGCTGGCGTGGGTCTGTTGGTCTGCGCGGC	2361
GCATCGCTGTTTATTGCGAACCCTTGGCGCGCTGGCGACGGGCTGCTGGTCTGCGCGGG	2457
CTGGCGCGGCTTCTTGGCTTTCGTTACGTCATGCGGCTGCAGAGCAACCCATGAAGGCCCT	2426
CTGGTGGCGCTTCTTGGCTACCGGTACATTTCCGCTCCGACGCAACCCATGAAGGCGCT	2522
GTACCTCTAACCACCAAGGAGCTCAAGAACCCACCAACCCGGA-CGCGTCCGGGAGGGCGAG	2490
GTACCGATCACCACGCGCGCTCAAGGACGAG-CCCGGGCGCAACCGCCGCGGCGAG	2583
GAGGGCGGCGACTTTGACGAGGCAAGCTAGCGAGGCGAGGAGATACGTTACATGGCCCT	2555
GAAGAGGAGGAGTTTACGCGGCAAACTGGAGAGGCGCGGAGATGATCAAGTATATGCTGCT	2648
GGTGTGCGCATGAGCGCAGCAACCAAGGCAAGGAAGGGCACGAGCG---CGTGTCTCA	2617
CGTGTGCGCGTGCAGCGGCAAGAGCACAAGCGAAGGAGAGCAACAGGGCGCGCGCTGCTGG	2713
GCGCCAAGGTACCCGACATGGTCAT-----GCGCA-----AGCGCGCAACCAACT	2665
-----CGACCGCGCTGACGACGCTCGCGCTTCGCGGCGAGCGCGCGGAG-----T	2761
ACACCAAGTTCCCAACAAAGACGG-----TGACGCGGACGAGGACGCTGTGACGGGG	2720
ACCAGCAGTTCGATGGCGACGTCGGGGGGGCTAGAGGCTA-----TGAT-GGGC	2814
POLY(A)	
GGTTTGTGTAATAAAACCAACGGGTGT-----TAAACCGCATGCGCATCTT-----	2768
AGTT-----CGGGTGAATAATAAATTTGCGCGAATCTTATTTAAGTGCA	2861
POLY(A)	
-----TTGGTTTTTTTGT-----	2783
CACCGTGTATTGCGGCTGTTGTTTTCTTGGAGCGGGACGTGCGCGAGCTCGGCCGAT	2926
---GGTCA-----GCCCTTTGTGTGTGTGGGAAGAAAGAAAAAG	2821
TAGGGTTTCGGCGCCACCGGGCACGGCAGGCGCCCTTACTTATGTTTGGCGCGGTGGCTCC	2991
GAACACATAAATC-----CCCCGGGTGTCGCGGCTGTTCTCTCTTTC	2866
GG-CACCGGT-CTCTGTGGCCTCCCGCCGCTTTCGCTTATTTGTCACAGCTG	3043
Pvu II	

FIG. 1. DNA sequences of the BHV-1 gI gene and HSV-1 gB gene. Nucleotide positions are numbered with reference to the translation initiation codons for each gene. The method of Wilbur and Lipman (33) was used to generate the alignment presented here. Identical bases are indicated by a colon. Translation initiation (ATG) and termination (TGA) codons are underlined. All other sequence features are underlined and labeled as they are referred to in the text [GC, CAAT, TATA, and POLY(A)]. Terminal *HpaI* and *PvuII* restriction endonuclease recognition sites of the BHV-1 sequence are also underlined and labeled.

Collection. Plaque-purified virus was propagated in Madin-Darby bovine kidney cells as described previously (14).

Bacterial strains. *Escherichia coli* DH5 frozen competent cells (Bethesda Research Laboratories, Gaithersburg, Md.) were used for all transformations with bacteriophage M13 replicative form (RF) DNA. *E. coli* JM101 was used to propagate bacteriophage M13 for the preparation of single- and double-stranded M13 DNA.

Cloning into M13 sequencing vectors. A 3.6-kb *HpaI-KpnI* restriction fragment capable of rescuing BHV-1 expressing antigenically altered gI was described previously (14) and was incorrectly designated a 3.8-kb fragment on the basis of electrophoretic mobility in agarose gels. A *Sall-KpnI* restriction fragment 48 nucleotides larger than the *HpaI-KpnI* fragment (the *Sall* site is 48 nucleotides outside of the *HpaI* site) was cloned into bacteriophage M13mp19 RF DNA which had been digested with both *Sall* and *KpnI* restriction endonucleases. The resultant recombinant bacteriophage M13mp19 was designated clone A1. The same *Sall-KpnI* restriction fragment was also cloned into bacteriophage M13mp18 RF DNA which had been digested with *SmaI* restriction endonuclease (yielding blunt ends). To accomplish this, the *Sall* and *KpnI* restriction endonuclease-generated ends of the 3.6-kb fragment were made blunt with mung bean nuclease. This blunt-ended fragment was then cloned into *SmaI*-digested, bacterial alkaline phosphatase-treated M13mp18 RF DNA. Two recombinant clones representing the two insert orientations were chosen and designated B1 (insert orientation the same as that of clone A1 with respect to the universal primer annealing site) and B51.

Sequencing strategy. The nucleotide sequence of the recombinant M13mp18 clone B51 insert was determined from a nested set of 16 individual clones generated by progressive, unidirectional digestion of insert DNA (starting from what was the *Sall* site before blunt ending and digesting toward the *KpnI* site). This was done by using the exonuclease III deletion method of Henikoff (10) modified by substituting mung bean nuclease for S1 nuclease. The sequence of the clone A1 recombinant M13mp19 insert was determined from the full insert by using synthetic oligodeoxynucleotides (17-mers) as sequencing primers. Primers were prepared on the basis of the known sequence generated from sequencing the B51 insert. Oligodeoxynucleotides were synthesized by using an Applied Biosystems model 380A automated DNA synthesizer which uses β -cyanoethyl phosphoramidite chemistry.

Determination of DNA nucleotide sequence. Sequencing was performed by the dideoxynucleotide chain termination method of Sanger et al. (26). The analog 7-deaza-2'-deoxy-GTP (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) was substituted for dGTP in all sequencing reactions to minimize band compression resulting from the high G+C content of the DNA. Deoxyadenosine 5'-[α - 35 S]thiotriphosphate (650 Ci/mmol) (Amersham Corp., Arlington Heights, Ill.) was used to radioactively label sequencing reaction products. Electrophoretically separated sequencing reaction products were visualized by autoradiography of the dried sequencing gel by using Kodak XAR film exposed at -70°C . Computer-generated hydropathic plots and similarity matrix analyses were done by using the IBI DNA/Protein Sequence Analysis System (International Biotechnologies, Inc., New Haven, Conn.). Nucleotide and amino acid sequence alignments were generated by using the ALIGN and AALIGN software programs, respectively (DNASTAR, Inc., Madison, Wis.).

RESULTS

Nucleotide sequence of BHV-1 gI and similarity to HSV-1 gB. Figure 1 shows the BHV-1 gI nucleotide sequence (lower line) aligned with the published sequence of HSV-1 (KOS) gB (upper line; 3). Gaps have been introduced in the sequences of maximize overall similarity. On the basis of the alignment presented in Fig. 1, the similarity between the two sequences is 56.3%, although several regions of alignment exhibit local similarities much higher than this.

Within the BHV-1 nucleotide sequence presented in Fig. 1, a single open reading frame of sufficient length to encode a protein the size of gI was identified. From the first ATG codon (position +1 to +3 of the gI sequence in Fig. 1), this open reading frame is 2,796 base pairs and would encode a primary translation product of 932 amino acids. The predicted 904-amino-acid gB primary translation product is similar in size. The ATG triplet at position +1 to +3 of the gI nucleotide sequence is likely to be the translation initiation codon of the gI reading frame for several reasons. This particular ATG is the first such codon appearing in the gI open reading frame and is the first ATG in any reading frame downstream of the proposed mRNA transcriptional start site (discussed below). Moreover, with a purine (G) at the -3 position and a G residue at the +4 position, this ATG is couched in a sequence arrangement which is favorable for translation initiation according to Kozak (12).

Potential promoter elements associated with mRNA transcription by the cellular enzyme RNA polymerase II are present upstream from the proposed gI translation initiation codon (Fig. 1). The sequence TATAT located at position -149 to -145 may function as a TATA box element. It should be noted that this sequence aligns with the TATA box of the HSV-1 gB promoter and is positioned within a 24-nucleotide region of which 22 nucleotides are identical between the two sequences (Fig. 1). Transcription by RNA polymerase II usually initiates with a purine some 19 to 27 nucleotides downstream of the TATA box (15). This would localize the 5' end of the gI transcript to the -126 to -118 region. Upstream from the TATA box at position -184 to -175 is a decanucleotide which matches the consensus binding sequence for the cellular transcription factor Sp1 (GC box) at 9 of 10 positions (2). Still further upstream at position -244 to -227 are two partially overlapping decanucleotides which match the consensus Sp1 binding sequence at 8 of 10 positions. Although a potential CAAT box has been identified within the gB promoter (21), we find no sequence upstream of the putative gI TATA box which resembles the CAAT box consensus sequence (15).

Downstream of the translation termination codon (TGA; position 2797 to 2799) are two sequence elements associated with mRNA cleavage and polyadenylation sites in eucaryotic mRNA. The first element, AATAAA, is highly conserved at cleavage and polyadenylation sites in eucaryotic mRNA (22). The second element, located roughly 30 to 60 nucleotides downstream of the AATAAA, is a GT-rich region. Such GT-rich regions are similarly associated with many known mRNA cleavage and polyadenylation sites and reportedly affect the efficiency of processing at those sites (1, 18, 19).

Predicted amino acid sequence and secondary structure of BHV-1 gI. The amino acid sequence of the predicted gI primary translation product, aligned with that of HSV-1 gB, is presented in Fig. 2. Gaps have been inserted into the sequences to maximize the amino acid similarity between

the two proteins. The alignment presented in Fig. 2 yields an overall similarity of 45.9%.

Hydropathic analysis of the gI amino acid sequence (presented in Fig. 3B) demonstrates the presence of two prominent hydrophobic peaks. The first peak is very near the amino terminus of the protein, while the second peak is near the carboxyl terminus. This pattern closely matches that of HSV-1 gB (Fig. 3A) and is consistent with a membrane

gB	MHQGPSWGRWFVVMALLGLTLGLVLSAASPSSPGTPGVAAATQAA	47
gI	MAARGGAERAGAGDGRGRQRRLRPRGVRLAALRGPAAGGAGARAALAAALLWATWALL	60
	NGGPATPAPPALGAAPTGDPKPKKPKKNPTPPRAGDNATVAAGHATLREHLRDIKAENTDAN	112
	LAAPAAGRATTTPAPPPE-EAASPAPPASPPGPDGDAASPDNDVRAALRLQAAGENS	124
	signal ↓	
	FVVCPPPTGATVVQFEQPRRCPTREGRQNYTEGIAVFKENIAPYKFKATMYKDVTSQVWFGH	177
	FFVCPSPSGATVRLAPARPCPEYGLGRNYTEGIVGIYKENIAPYTFKAYI-YKNVIVTTTWAGS	188
	RYSQFMGIFEDRAPVPFEEIDKINAKGVCSTAKYVRNNLETTAFHRDDHETDMLKPANAATR	242
	TYAATINQYTDVRVPMGEITDLVOKKWRCLSKAEYLRSGRVAFDRDDDPWEAPLKPRLSAP	253
	TSRGMHTDLKYNPSRVEAFHRYGTTVNCIVEEDARSVPYDEFVLATGDFVMSPFYGYREGS	307
	GVRGMHTDDVYALGSGAGLYRTGTSVNCIVEEARSVPYDSFALSTGDIIMSPFYGLREGA	318
	HTEHTTYAADRFKQVDGFYARDLTTKARATPTTRNLLTPKFTVANDWVPKRPSVTCTMKQEV	372
	HREHTSYSPERFQIEGYKRDMAIGRRLKEPVSRNFLRTQHVTVANDWVPKRKNVCSLAKWREA	383
	DEMLRSEYGGSFSSDAISTFTTNLTLEYPLSRVDLGDGICDKARDAMDRIFFARRYNATHIKVG	437
	DEMLRDESGRNFRFTARSLSATFVSDSHTFALQNVPLSDVCIEEAEEAVERVRYNGTHVLGS	448
	QPQYLANGGFLIAYQPLLSNTLAELYVREHLREQS-----RKPPNPT	480
	SLETYLARGGFVAFRPMLSNELAKLYLQELARSNGTLEGLFAAAAPKPGRRARRAAPSAPGGP	513
	PPPPGASANASVE-RIKTTSSIEFARLQFTYHNIQRHVNDMLGRVIAIANCELQNHETLWNEARK	544
	GAANGPAGDGDAGGRVTVSSAEFAALQFTYDHIQHVNTMFSRLATSWCLLQNKERALEWAAK	578
	LNPNAIASVTGRRVSARMLGDMVAVSTCVPAADNVIVQNSMRISRRPGACYSRPLVSFRYEDQ	609
	LNPASAAASALDRRAAARMLGDMAVTYCHELGEGRVFIENSMRPG--GVCYSRPPVSAFAGNE	641
	GPLVEGQLGENNELRLTRDAIEPCTVGHRRYFTFGGGVYVEEYAYSHQLSRADITTVSTFDLN	674
	SEPVEGQLGEDNELLPGRELVEPCTANHRYFRFGADYVYENYAVVRRVPLAELEVISTFDLN	706
	ITMLEDHEFVPLEVYTRHEIKDGLLDYTEVQRRNLHDLRFADITVIHADANAAMFAGLGAF	739
	LTVLEDREFLPLEVYTRAEALDGLDYSEIQRNLHELRFYDIDRVVKTGDMAIMRGLANFF	771
	transmembrane region	
	EGMGDLGRAVGKVMGIVGGVSAVSGVSSFMSPFGALAVGLLVLAGLAAFAFRYVMRLQSN	804
	QGLGAVGQAVGTVVLGAAGAALSTVSGIASFIANPFGALATGLLVLAGLVAALAYRYSRLRSN	836
	PMKALYPLTTKELKNPTNDASGEEGGDFDEAKLAEAREMIRYMALVSAMERTEHKAKKKGTS	869
	PMKALYPIITRALKD-DARGATAPGEEEEFDAAKLEQAREMIKYSLSVAVERQEHKAKKSNKG	900
	A-LLSAKVTDMMKRNRNTNYQVPMKGDGDAEDDL	904
	GPLLATRLTQLALRRRAPPEYQQLPMADVGGA	932

FIG. 2. Predicted amino acid sequence of BHV-1 gI and HSV-1 gB. Identical amino acids are indicated by a colon. Potential N-linked glycosylation sites are underlined. Predicted signal peptidase cleavage sites and transmembrane region boundaries are indicated by arrows and labeled. The alignment shown was generated independently of the nucleotide sequence alignment and takes into consideration similar, as well as identical, amino acid residues, although only identical residues are indicated.

FIG. 3. Hydropathic analysis of the BHV-1 gI and HSV-1 gB predicted amino acid sequences. The graphic hydropathic analyses shown were generated by the method of Kyte and Doolittle (13) with a moving window of 9 amino acids. The x axis represents the average hydropathicity value. Hydrophobic regions are indicated by shaded areas above the x axis while hydrophilic regions are indicated by shaded areas below the x axis. Amino acid residues are numbered as in Fig. 2. Numbering is from left to right and appears below the individual hydropathic analyses. Plots are compressed 4 times.

orientation like that proposed for gB. Glycoprotein B appears to be oriented such that the amino terminus and the bulk of the polypeptide backbone face the extracellular environment while a relatively small carboxyl-terminal portion faces the cytoplasm or virion envelope interior (5). The amino-terminal hydrophobic peak of gI probably represents the signal sequence of the protein, directing its synthesis to the rough endoplasmic reticulum. Analysis of the amino acid sequence in this region of gI by the method of von Heijne (32) confirms its signal sequence-like character and predicts a potential signal peptidase cleavage site after Gly 67. Thus, after cleavage of the putative signal sequence, the gI polypeptide would contain 865 amino acids.

(i) **The unusual gI amino-terminal leader.** Unlike that of gB, the predicted gI signal sequence is preceded by a region of predominantly hydrophilic amino acids which is evident in the hydropathic analysis presented in Fig. 3. An almost identical situation has recently been reported for the gB homolog of pseudorabies virus (PrV), gII (23). These authors recognized that a portion of an open reading frame largely upstream of gII overlaps the first 44 codons of the gII coding

region. This upstream open reading frame was shown to exhibit both nucleotide and amino acid sequence similarity to the HSV-1 ICP18.5 coding region which occurs upstream of, but does not overlap, the gB coding region (20). Upon further inspection of the gI nucleotide sequence, we found that, as in PrV, part of an upstream open reading frame showing extensive similarity to the HSV-1 ICP18.5 coding region overlaps the gI coding region by roughly 47 codons (depicted in Fig. 4A). As suggested by Robbins et al. (23) for the PrV gII leader, the unusually long BHV-1 gI leader may reflect the constraints imposed upon the sequence to accommodate both an upstream coding region and a functional signal sequence.

(ii) **The transmembrane domain.** The broad hydrophobic peak occurring near the carboxyl terminus of gI probably functions as a membrane-spanning region. This peak represents roughly 69 amino acids (residues 759 to 827) and is similar in length to the proposed gB transmembrane region. Given the length and predicted secondary structure characteristics of the gB transmembrane region, it has been suggested that it spans the membrane three times as contiguous,

FIG. 4. Arrangement of the BHV-1 gI coding region and upstream open reading frame compared with the arrangement of HSV-1 homologous coding regions. (A) Arrows represent open reading frames and indicate the direction of translation. ORF, The BHV-1 open reading frame upstream of, and partially overlapping, the gI coding region; all other coding regions are labeled as they are referred to in the text. The dashed region of ORF represents the predicted (not yet sequenced) portion of that open reading frame. This diagram is not drawn to scale. (B) Open bars represent open reading frames. Shaded areas represent gaps in the respective nucleotide sequences according to the alignment presented in Fig. 1. Relevant translation initiation (ATG) and termination (TAG) codons are indicated. Reading frames are labeled as in panel A. The BHV-1 arrangement is shown with two separate bars for clarity. This diagram is drawn to scale. Note that a counterpart to the region of the HSV-1 sequence which includes the ICP18.5 translation terminator and the gB translation initiator is apparently not present in BHV-1. The BHV-1 gI coding region begins in an area of the sequence showing similarity to a portion of the HSV-1 ICP18.5 coding region, while the upstream open reading frame terminates in an area exhibiting sequence similarity to a portion of the gB coding region.

antiparallel alpha-helices (21). It seems possible that the BHV-1 gI transmembrane region could behave similarly since its length and predicted secondary structure characteristics (data not shown) are similar to those of gB (the gI and gB amino acid sequences exhibit 59% similarity through this region).

(iii) **The extracellular domain.** It has been shown previously that gI contains N-linked oligosaccharide moieties (29). Analysis of the gI amino acid sequence reveals the presence of six potential N-linked glycosylation sites (Asn - X - Ser/Thr; X $\neq$ Pro) in the putative extracellular domain. HSV-1 gB also contains six potential N-linked glycosylation sites, although the positions of only four sites are colinear with those of gI by the alignment in Fig. 2. It should be noted that the sixth potential N-linked glycosylation site of gB, which aligns with the sixth site of gI, is apparently not used (S. Person, personal communication). Both BHV-1 gI and HSV-1 gB contain 10 Cys residues, all of which are located in the extracellular domains of the glycoproteins. The amino acid sequence alignment in Fig. 2 demonstrates the remarkable conservation of the positions of all 10 Cys residues between the two glycoproteins.

The amino acid change responsible for the rapid host cell entry phenotype of the HSV-1 HFEM isolate, *tsB5*, has been localized to the gB extracellular domain (4). The site of this

substitution (gB residue 553) occurs in a region of the gB sequence exhibiting relatively high local similarity with gI (65% similarity over 72 amino acids; gB residues 499 to 570). Interestingly, gI amino acid 587, which is colinear with gB residue 553, is an alanine which corresponds to the amino acid change in gB responsible for the *tsB5* rapid-entry phenotype.

One additional noteworthy feature of the gI extracellular domain is the presence of a 22-amino-acid region (residues 485 to 506) for which there is essentially no counterpart in gB. A similar situation has been noted for the gB homolog of PrV, gII. PrV gII, like BHV-1 gI (but not gB), exists in its mature form largely as cleaved, disulfide-linked molecules, and it has been suggested that the cleavage site of PrV gII lies within this region of nonsimilarity with gB (23). These authors have proposed that the sequence Arg-Arg-Ala-Arg-Arg, which occurs within the PrV gII region of nonsimilarity with gB, may be the processing site for that protein. This same 5-amino-acid sequence also uniquely occurs within the above-noted region of BHV-1 gI which lacks a gB counterpart.

(iv) **The cytoplasmic domain.** The predicted 105-amino-acid cytoplasmic domain of gI is largely hydrophilic, as indicated by the hydropathic analysis (Fig. 3B), with charged amino acids accounting for nearly one-third of the residues in this region. These general characteristics are common to the cytoplasmic domains of several other transmembrane glycoproteins, both viral and cellular (25), and may thus satisfy a basic requirement of the host cell machinery (perhaps necessary to arrest transfer across the endoplasmic reticulum). The similarity between the cytoplasmic domains of gI and gB is roughly 51% overall, with two regions of 19 and 23 amino acids exhibiting local similarities of 79 and 83%, respectively. The amino acid change responsible for the Syn (syncytium-forming) phenotype of the HSV-1 isolate, *tsB5*, is an Arg to His at amino acid 858 of gB. This position lies in the cytoplasmic domain within the 23-amino-acid region of 83% local similarity with gI mentioned above. The relatively high levels of local similarity between regions of the gI and gB cytoplasmic domains may reflect a common role for this domain beyond a simple anchor-stop transfer function. Some obvious possibilities include roles in recognition by, or interaction with, cellular proteins or other viral proteins during virion assembly. The relatively large sizes of the gI and gB cytoplasmic domains are consistent with such possibilities.

DISCUSSION

Previously, this laboratory reported the genomic map position of the BHV-1 gene encoding the gI glycoprotein (14). BHV-1 expressing an antigenically altered gI was rescued by a 3.6-kb *HpaI-KpnI* restriction fragment from the long unique region of the wild-type BHV-1 genome (0.405 to 0.432 map units). Additionally, nick-translated HSV-1 gB DNA was used to probe Southern blots of restriction fragments representing the entire BHV-1 genome. The HSV-1 gB probe uniquely hybridized to the same 3.6-kb *HpaI-KpnI* fragment. This result demonstrated a degree of DNA sequence similarity between these two herpesvirus glycoprotein genes. In this report, we present the nucleotide sequence and deduced translation of the BHV-1 gI gene. Comparison with the HSV-1 gB gene reveals an overall similarity of 56.3% at the nucleotide level and 45.9% at the amino acid level.

Several sequence elements potentially involved in mRNA synthesis and processing flank the gI coding region. Potential

FIG. 5. Similarity matrix analysis of the BHV-1 gI amino acid sequence versus the homologous glycoproteins of four other herpesviruses. In each plot, the horizontal axis is the BHV-1 amino acid sequence numbered from left to right as in Fig. 2. The vertical axes are the amino acid sequences of the gI homologs of the indicated viruses. Vertical axes are numbered from top to bottom. Points are plotted when a minimum of 9 amino acids within a moving window of 21 are identical. Uninterrupted diagonal lines from upper left to lower right indicate similarity over larger areas. All axes are compressed 10 times. Vertical axes represent PrV gII (23) (A), HSV-1 gB (3) (B), varicella-zoster virus gPII (7) (C), and human cytomegalovirus gB (6) (D).

RNA polymerase II promoter elements present upstream of the gI coding region include a TATA box and several binding sites for the cellular transcription factor Sp1 (GC boxes). The sequence AATAAA, followed by a GT-rich region some 30 nucleotides further downstream, is present just 3' to the gI translation termination codon and probably constitutes a mRNA cleavage and polyadenylation signal.

In addition to the amino acid sequence similarity of 45.9%, the BHV-1 gI and HSV-1 gB proteins share several other features. The hydropathic analyses are very similar in shape and exhibit two prominent peaks of hydrophobicity which likely represent the signal sequence (at the amino terminus and absent from the mature protein) and the transmembrane domain (near the carboxyl terminus). Both proteins contain 6 potential N-linked oligosaccharide addition sites in their predicted external domains, four of which are colinear, as

well as 10 highly conserved Cys residues. Several local regions of high similarity between the two amino acid sequences are present, including those regions which contain the two known amino acid changes within gB of the HSV-1 isolate, *tsB5*. These changes are responsible, separately, for the rapid rate of virus entry phenotype (gB residue 553) and the syncytium-forming phenotype (gB residue 858) of this HSV-1 mutant (4). The extensive similarity exhibited suggests that these two glycoproteins may share important structural and functional properties.

Two regions of little or no similarity between the gI and gB amino acid sequences are apparent from the alignment presented in Fig. 2. These regions correspond to a central region of the gI sequence which may contain a proteolytic processing site not present in gB and to the amino-terminal leaders of the glycoproteins. One might expect that the

amino-terminal leaders would be free to mutate as long as the requirements of a functional signal sequence are met, since this portion of the translated polypeptide is absent from the mature glycoprotein molecule. Consistent with this, the gB and gI leaders are quite different both in length and amino acid sequence. Each leader, however, contains a sequence of roughly 20 amino acids which is strongly hydrophobic and probably represents a signal sequence on the basis of its amino-terminal location. Recent evidence suggests that, aside from the signal peptidase cleavage site, perhaps the sole requirement of a functional signal sequence is an appropriately positioned hydrophobic core (11).

The recently reported nucleotide sequence of the PrV gII glycoprotein gene (also a homolog of HSV-1 gB) predicts an amino-terminal leader of 53 amino acids with a hydrophilic region preceding the proposed signal sequence much like the arrangement seen in BHV-1 gI (23). The character of the PrV gII leader has been attributed to an apparent overlap between the coding regions of the gII leader and a potential upstream gene. The sequenced portion of this potential upstream gene exhibits extensive similarity at both the nucleotide and amino acid levels to the HSV-1 ICP18.5 gene (20). The coding region of the HSV-1 ICP18.5 gene is present upstream of, but does not overlap, the gB coding region. We have found that, as in PrV, a portion of an upstream open reading frame exhibiting extensive nucleotide and amino acid sequence similarity (58.9% similarity over 175 amino acids) to the comparable region of the HSV-1 ICP18.5 gene exists and overlaps the first 47 codons of the predicted gI coding region.

Figure 4B depicts the end of the ICP18.5 coding region and the beginning of the gB coding region as they occur in HSV-1 aligned with the comparable coding regions of BHV-1. In this diagram, coding regions are shown containing gaps corresponding to those present in the nucleotide sequence alignment of Fig. 1. Evident from this diagram is a large gap in the BHV-1 sequence encompassing the HSV-1 ICP18.5 termination codon as well as the gB initiation codon. The BHV-1 upstream open reading frame thus extends to the next in-frame termination codon (TAG; BHV-1 position +143). An upstream, nonhomologous initiation codon for gI is used. It is interesting that the BHV-1 initiation codon (ATG) is directly opposite a CTG in the HSV-1 sequence in the Fig. 1 alignment. Consequently, a single base change in the HSV-1 sequence could generate an ATG upstream of the actual gB initiation codon and within the ICP18.5 coding region.

The gap located centrally in the gB amino acid sequence in the alignment presented in Fig. 2 may represent the lack of a proteolytic processing site which is present in gI and is responsible for cleavage into the 74- and 55-kDa forms. A similar situation has been reported for the PrV gII glycoprotein, which is also proteolytically processed to 68- and 58-kDa forms which remain disulfide linked (9).

Although this report has focused primarily on the similarity between BHV-1 gI and HSV-1 gB, it should be noted that the reported gB homologs of several other herpesviruses also share detectable similarity to BHV-1 gI. The matrix analyses presented in Fig. 5 illustrate the similarity at the amino acid level between BHV-1 gI and PrV gII, HSV-1 gB, varicella-zoster virus gpII, and human cytomegalovirus gB (plots A through D, respectively). Each point plotted represents a minimum similarity of 42% over 21 amino acid residues. Uninterrupted diagonals from upper left to lower right represent longer stretches of similarity. That gB is essential to HSV-1 infectivity has been demonstrated by the

ability to isolate conditional lethal gB mutants (16, 27). The observed amino acid sequence conservation between the gB homologs of various herpesviruses may reflect a common, essential function of these molecules.

ACKNOWLEDGMENTS

This work was supported by U.S. Department of Agriculture (USDA) Grant 86-CRCR-1-2221, USDA Grant 83-CRSR-2-2237, USDA section 1433 Formula Funds, and Commonwealth of Pennsylvania, Department of Agriculture Grant ME-4606. J.C.W. was the recipient of a Merck Fellowship.

We thank M. J. Mitchell of the University of Pennsylvania Cancer Center DNA Synthesis Service for synthesis of oligodeoxynucleotides.

LITERATURE CITED

1. Birnstiel, M. L., M. Busslinger, and K. Strub. 1985. Transcription termination and 3' processing: the end is in site! *Cell* **41**: 349-359.
2. Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* **234**:47-52.
3. Bzik, D. J., B. A. Fox, N. A. DeLuca, and S. Person. 1984. Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex virus type 1. *Virology* **133**:301-314.
4. Bzik, D. J., B. A. Fox, N. A. DeLuca, and S. Person. 1984. Nucleotide sequence of a region of the herpes simplex virus type 1 gB glycoprotein gene: mutations affecting rate of virus entry and cell fusion. *Virology* **137**:185-190.
5. Claesson-Welsh, L., and P. G. Spear. 1987. Amino-terminal sequence, synthesis, and membrane insertion of glycoprotein B of herpes simplex virus type 1. *J. Virol.* **61**:1-7.
6. Cranage, M. P., T. Kouzarides, A. T. Bankier, S. Satchwell, K. Weston, P. Tomlinson, B. Barrell, H. Hart, S. E. Bell, A. C. Minson, and G. L. Smith. 1986. Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. *EMBO J.* **5**:3057-3063.
7. Davidson, A. J., and J. E. Scott. 1986. The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* **67**:1759-1816.
8. Gibbs, E. P. J., and M. M. Rweyemamu. 1977. Bovine herpesviruses. I. Bovine herpesvirus 1. *Vet. Bull.* **47**:317-343.
9. Hampl, H., T. Ben-Porat, L. Ehrlicher, K.-O. Habermehl, and A. S. Kaplan. 1984. Characterization of the envelope proteins of pseudorabies virus. *J. Virol.* **52**:583-590.
10. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**:351-359.
11. Kaiser, C. A., D. Preuss, P. Grisafi, and D. Botstein. 1987. Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science* **235**:312-317.
12. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283-292.
13. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**:105-132.
14. Lawrence, W. C., R. C. D'Urso, C. A. Kundel, J. C. Whitbeck, and L. J. Bello. 1986. Map location of the gene for a 130,000-dalton glycoprotein of bovine herpesvirus 1. *J. Virol.* **60**:405-414.
15. Lewin, B. 1987. *Genes*. John Wiley & Sons, Inc., New York.
16. Little, S. P., J. T. Jofre, R. J. Courtney, and P. A. Schaffer. 1981. A virion-associated glycoprotein essential for infectivity of herpes simplex virus type 1. *Virology* **115**:149-160.
17. Marshall, R. L., L. L. Rodriguez, and G. J. Letchworth III. 1986. Characterization of envelope proteins of infectious bovine rhinotracheitis virus (bovine herpesvirus 1) by biochemical and immunological methods. *J. Virol.* **57**:745-753.
18. McDevitt, M. A., R. P. Hart, W. W. Wong, and J. R. Nevins. 1986. Sequences capable of restoring poly(A) site function define two distinct downstream elements. *EMBO J.* **5**:2907-

- 2913.
19. **McLauchlan, J., D. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. *Nucleic Acids Res.* **13**:1347-1368.
20. **Pellett, P. E., F. J. Jenkins, M. Ackermann, M. Sarmiento, and B. Roizman.** 1986. Transcription initiation sites and nucleotide sequence of a herpes simplex virus 1 gene conserved in the Epstein-Barr virus genome and reported to affect the transport of viral glycoproteins. *J. Virol.* **60**:1134-1140.
21. **Pellett, P. E., K. G. Kousoulas, L. Pereira, and B. Roizman.** 1985. Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene: primary sequence and predicted protein structure of the wild type and of monoclonal antibody-resistant mutants. *J. Virol.* **53**:243-253.
22. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3' non-coding region sequences in eukaryotic messenger RNA. *Nature (London)* **263**:211-214.
23. **Robbins, A. K., D. J. Dorney, M. W. Wathen, M. E. Whealy, C. Gold, R. J. Watson, L. E. Holland, S. D. Weed, M. Levine, J. C. Glorioso, and L. W. Enquist.** 1987. The pseudorabies virus gII gene is closely related to the gB glycoprotein gene of herpes simplex virus. *J. Virol.* **61**:2691-2701.
24. **Roizman, B., L. E. Carmichael, F. Deinhardt, G. de The, A. J. Nahmias, W. Plowright, F. Rapp, P. Sheldrich, M. Takahashi, and K. Wolf.** 1982. Herpesviridae: definition, provisional nomenclature and taxonomy. *Intervirology* **16**:201-217.
25. **Sabatini, D. D., G. Kreibich, T. Morimoto, and M. Adesnik.** 1982. Mechanisms for the incorporation of proteins in membranes and organelles. *J. Cell Biol.* **92**:1-22.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
27. **Sarminento, M., M. Haffey, and P. G. Spear.** 1979. Membrane proteins specified by herpes simplex viruses. III. Role of glycoprotein VP7(B₂) in virion infectivity. *J. Virol.* **29**:1149-1158.
28. **Spear, P. G.** 1985. Glycoproteins specified by herpes simplex viruses, p. 315-356. *In* B. Roizman (ed.), *The herpesviruses*. Plenum Publishing Corp., New York.
29. **van Drunen Littel-van den Hurk, S. and L. A. Babiuk.** 1985. Effect of tunicamycin and monensin on biosynthesis, transport and maturation of bovine herpesvirus type-1 glycoproteins. *Virology* **143**:104-118.
30. **van Drunen Littel-van den Hurk, S., and L. A. Babiuk.** 1986. Synthesis and processing of bovine herpesvirus 1 glycoproteins. *J. Virol.* **59**:401-410.
31. **van Drunen Littel-van den Hurk, S., J. V. van den Hurk, J. E. Gilchrist, V. Misra, and L. A. Babiuk.** 1984. Interactions of monoclonal antibodies and bovine herpesvirus type 1 (BHV-1) glycoproteins: characterization of their biochemical and immunological properties. *Virology* **135**:466-479.
32. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14**:4683-4690.
33. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA* **80**:726-730.