JOURNAL OF VIROLOGY, Nov. 1991, p. 5744–5750
0022-538X/91/115744-07$02.00/0

Vol. 65, No. 11

Structure and Genetic Variability of Envelope Glycoproteins of Two Antigenic Variants of Caprine Arthritis-Encephalitis Lentivirus

D. P. KNOWLES, JR.,[1]* W. P. CHEEVERS,[2] T. C. McGUIRE,[2] A. L. BRASSFIELD,[2] W. G. HARWOOD,[1] AND T. A. STEM[2]

USDA Agricultural Research Service Animal Disease Research Unit[1] and Department of Veterinary Microbiology and Pathology,[2] College of Veterinary Medicine, Washington State University, Pullman, Washington 99164-7040

Received 24 April 1991/Accepted 22 July 1991

To define the structure of the caprine arthritis-encephalitis virus (CAEV) *env* gene and characterize genetic changes which occur during antigenic variation, we sequenced the *env* genes of CAEV-63 and CAEV-Co, two antigenic variants of CAEV defined by serum neutralization. The deduced primary translation product of the CAEV *env* gene consists of a 60- to 80-amino-acid signal peptide followed by an amino-terminal surface protein (SU) and a carboxy-terminal transmembrane protein (TM) separated by an Arg-Lys-Lys-Arg cleavage site. The signal peptide cleavage site was verified by amino-terminal amino acid sequencing of native CAEV-63 SU. In addition, immunoprecipitation of [^{35}S]methionine-labeled CAEV-63 proteins by sera from goats immunized with recombinant vaccinia virus expressing the CAEV-63 *env* gene confirmed that antibodies induced by *env*-encoded recombinant proteins react specifically with native virion SU and TM. The *env* genes of CAEV-63 and CAEV-Co encode 28 conserved cysteines and 25 conserved potential N-linked glycosylation sites. Nucleotide sequence variability results in 62 amino acid changes and one deletion within the SU and 34 amino acid changes within the TM.

Caprine arthritis-encephalitis virus (CAEV) is a lentivirus which causes persistent infection and inflammation in synovial spaces (6, 11), resulting in recurrent swelling of carpal joints and progressive arthritis in clinically affected goats (6). Previous studies on the pathogenesis of CAEV-induced arthritis have provided impetus to analyze the genetic structure and variability of the CAEV envelope (*env*) gene. These studies have shown that the immune response of goats to CAEV envelope proteins is involved in the pathogenesis of arthritis (25, 27, 34).

The dominant antibody response of goats to CAEV infection is directed against the gp135 surface glycoprotein (SU) (25), and synovial fluid anti-SU antibody titers correlate specifically with the presence and severity of arthritis (27). A collateral experiment demonstrated that goats immunized with inactivated CAEV develop more rapid and more severe arthritis after virus challenge than nonimmunized controls (34), confirming that immune responses to viral antigen are involved in exacerbation of joint inflammation. In addition to dominant anti-SU antibody responses, some goats develop neutralizing antibody directed to type-specific SU epitopes (30), allowing identification of antigenic variants of CAEV from synovial fluid of arthritic joints (13, 36). In this regard, it has been demonstrated that (i) some antigenic variants produced by persistent virus replication are clonally expanded within carpal joints of CAEV-infected goats due to a replication advantage conferred by resistance to type-specific neutralizing antibody and (ii) recurrent emergence of successive neutralization-resistant variants occurs in association with progressive arthritis (7). Thus, virus replication and development of progressive inflammatory lesions are not arrested by the presence of naturally evolved neutralizing antibody, and replicating antigenic variants of CAEV may contribute to the progression of arthritis.

These observations indicate that at least two classes of CAEV SU epitopes are involved in the pathogenesis of progressive arthritis caused by this lentivirus. Specifically, we hypothesize that the progressive course of arthritis is sustained by anamnestic anti-SU immune responses to conserved immunopathogenic epitopes expressed by variants of CAEV which attain a replication advantage by *env* gene mutations that alter the specificity of SU neutralization epitopes. To pursue this hypothesis, studies were initiated to identify and characterize antigenic determinants of CAEV envelope proteins. In the present work, the structure and genetic variability of proteins encoded by the *env* genes of two antigenic variants, CAEV-63 and CAEV-Co (36), are described.

MATERIALS AND METHODS

Viruses and cells. CAEV-63 (11) and CAEV-Co (39) were propagated in goat synovial membrane (GSM) cells grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5 to 10% fetal bovine serum as described previously (26).

RNA isolation and cDNA cloning. Polyadenylated cytoplasmic RNA was isolated from CAEV-infected GSM cells by modifications of described methods (33) and purified by oligo(dT)-cellulose column chromatography. cDNA libraries were produced in a modified pUC19 vector and amplified in *Escherichia coli* DH5. pUC19 was modified by adding an *Nsi*I site to the polylinker sequence at the *Bam*HI site. The *Nsi*I site provided digested 3′ overhangs for efficient homopolymeric dCMP tailing of the vector and allowed excision of cDNA from recombinant plasmids by digestion of flanking *Sal*I and *Sma*I polylinker sites not present within the CAEV-63 *env* gene (48).

cDNA synthesis and dGMP homopolymeric tailing reactions have been described before (18, 37). First-strand cDNA synthesis was primed with a 20-mer oligonucleotide complementary to CAEV-63 U3 (Fig. 1). After annealing of multiple ratios of G-tailed cDNA and C-tailed vector, recom-

* Corresponding author.

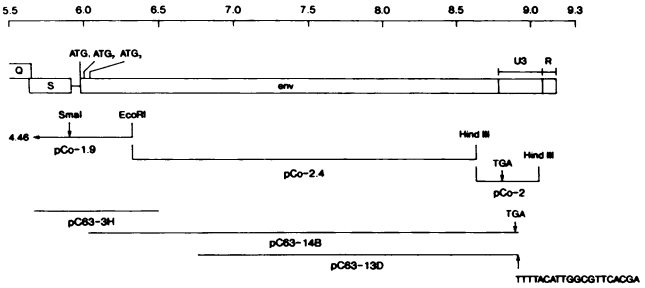

FIG. 1. Genetic map of 3′ half of CAEV genome, illustrating CAEV-Co and CAEV-63 molecular clones selected for nucleic acid sequencing. Derivation of clones is described in the text. Q and S have been previously defined (24, 51). The scale at the top is in kilobases.

binant plasmids were amplified in *E. coli* DH5 as described before (19) and screened by colony hybridization with the *Hin*dIII insert of pCo-1 labeled with [^{32}P]dCTP. pCo-1 is a pUC13 construct (46) containing 20 nucleotides of 3′ U3 and R-U5 regions of the long terminal repeat (LTR) and the colinear sequence of CAEV-Co proviral DNA extending to a *Hin*dIII site 155 nucleotides upstream of the *env* termination codon (20, 51). Plasmid DNA was isolated from positive clones by a plasmid minipreparation procedure (22), linearized with *Sal*I, and examined by agarose gel electrophoresis; recombinant plasmids with inserts of >0.5 kb were screened for the presence of CAEV-63 *env* gene restriction sites (48). On the basis of these criteria, CAEV-63 cDNA clones pC63-13D and pC63-14B (Fig. 1) were selected for sequencing. Construction of pC63-3H cDNA (Fig. 1) in C-tailed pUC9 was reported previously (24).

DNA sequencing and sequence analysis. CAEV-63 cDNA was recovered from pC63-13D by digestion with *Sal*I and *Kpn*I and subcloned for sequencing into pBluescript II SK$^+$ (pBSK$^+$) (Stratagene). pC63-14B was sequenced in the *Nsi*I-modified pUC19 cDNA cloning vector. The sequence of CAEV-63 open reading frame (ORF) S (*tat*) in pC63-3H was reported (24); the sequence of the 5′ terminus of the *env* ORF in this clone is reported here. Three CAEV-Co clones derived from pCo-1 or pCo-2 (46) were used to obtain the sequence of the *env* gene of this variant (Fig. 1). (i) The CAEV-Co clone designated pCo-2 is a pGEM construct containing the 3′-terminal 155 nucleotides of the *env* ORF and the 5′-terminal 267 nucleotides of U3 (46, 51). (ii) pCo-2.4 is a pUC19 subclone of pCo-1 extending from an *Eco*RI site within the *env* ORF to the *Hin*dIII site defining the junction with pCo-2. (iii) pCo-1.9 is a pBSK$^+$ subclone of pCo-1 extending from an *Eco*RI site within *pol* to the *Eco*RI site defining the junction with pCo-2.4 (46). The CAEV-Co *tat* sequence in pCo-1.9 has been reported (24); the sequence of the 5′ terminus of the *env* ORF in pCo-1.9 is reported here.

The cDNA inserts of pC63-13D and pC63-14B and the *Eco*RI insert of pCo-2.4 were sequenced from deletion libraries of these plasmids produced by using an exonuclease III-mung bean nuclease protocol (Stratagene) previously used to sequence pCo-1.9 (24). pC63-13D was digested at the pBSK$^+$ *Eco*RI and *Sac*I sites and deleted from the *Eco*RI site with exonuclease III. pC63-14B was digested at the pUC19 *Sma*I and *Sac*I sites and deleted from the *Sma*I site. pCo-2.4 was digested at the pUC19 *Sal*I and *Sph*I sites and deleted from the *Sal*I site. Deleted molecules were blunt-ended with mung bean nuclease, circularized with T4 ligase, and amplified in *E. coli* XL1-Blue. Deletion clones varying by approximate 200-bp increments were identified by agarose gel electrophoresis. Dideoxy sequencing of deletion clones was done with a Sequenase kit (U.S. Biochemical) with [^{35}S]thio-dATP and M13/pUC forward and reverse primers (Boehringer Mannheim) or T3 and T7 pBSK$^+$ primers (Stratagene). Sequence gaps were completed with primers synthesized with a 381A DNA synthesizer (Applied Biosystems), and compressions were resolved by using *Taq* polymerase (Stratagene). The sequence of overlapping deletion clones was aligned by computer (12). pCo-2 was sequenced directly by using primers complementary to the SP6 and T7 promoter regions of pGEM.

Nucleotide sequences were translated and aligned with the MAP and GAP-2 sequence analysis programs (12). Amino acid sequences were compared by the FASTA program (16, 43).

RP-HPLC purification and amino-terminal sequencing of CAEV-63 gp135 SU. The codon specifying the amino terminus of gp135 SU was identified within the CAEV-63 *env* gene from the amino-terminal sequence of native gp135 SU. CAEV-63 released into the medium of infected GSM cells was collected by differential centrifugation (8) and suspended in DMEM. The virus was disrupted in 2 volumes of 6 M guanidinium hydrochloride (pH 3.1), and viral proteins were separated by reverse-phase high-pressure liquid chromatography (RP-HPLC) (3). RP-HPLC fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (17) and by Western immunoblots probed with immune goat serum and ^{125}I-labeled recombinant protein G (21). Fractions containing gp135 were combined, lyophilized, dissolved in SDS-PAGE sample buffer (25), and heated at 55°C for 10 min (23). Proteins were then subjected to SDS-PAGE, modified as described before (23), to minimize destruction of amino acid side chains and to prevent amino-terminal blockage by gel contaminants. Proteins were electrophoretically transferred to a polyvinylidene difluoride membrane (Immobilon-P; Millipore), stained with 0.1% Coomassie blue, and destained with 45% methanol–10% acetic acid. The portion of the membrane with bound gp135 was cut out and used to derive the amino-terminal sequence of gp135 by previously described methods (10) with an amino acid sequencer (Applied Biosystems model 470A).

Labeling of virus and immunoprecipitation. CAEV was labeled with [^{35}S]methionine and isolated from the culture medium of infected GSM cells by differential centrifugation (8). The methods for preparation of virus lysates, immunoprecipitation reactions, SDS-PAGE, and fluorography have been described (17), except that immune complexes were precipitated with recombinant protein G (ImmuBind AbSorbent; Genex).

Immunization of goats with recombinant vaccinia viruses and recombinant CAEV gp135 SU. The construction of recombinant vaccinia virus expressing the *env* gene of CAEV-63 has been described before (30). Briefly, the CAEV-63 *env* insert from cDNA pC63-14B was cloned into the *Sma*I site of the vaccinia virus expression vector pSC-11 (5), and recombinant vaccinia virus, designated rWR-63, was derived by homologous recombination with WR vaccinia virus (31, 32). A control vaccinia virus (rWR-SC11) was also derived from pSC-11. rWR-63 and rWR-SC11, used for vaccination, were propagated in HeLa S3 spinner cultures and purified by sucrose density gradient centrifugation (32). The viruses were suspended at a titer of 10^8 PFU/0.1 ml in 1 mM Tris, pH 9. Recombinant antigen for immunization was

prepared from the culture medium of GSM cells infected with rWR-63 or rWR-SC11 as described before (30). Goats 90G65 and 90G66 were immunized with rWR-63, and goats 90G09 and 90G62 were immunized with rWR-SC11. Each goat received three intradermal vaccinations at two sites on the neck with 10^8 PFU per site. The second and third immunizations were given at 1- and 3-month intervals. Each goat was given a final intramuscular immunization with one roller bottle equivalent of medium from GSM cells infected with either rWR-63 or rWR-SC11 in Freund's complete adjuvant.

Animals, virus infection, and sera. All goats were purebred Saanens obtained from CAEV-free dams. Sera 27 and 79 were from goats 83G27 and 85G79, infected by feeding 100 ml per day of colostrum containing 10^6 $TCID_{50}$ (50% tissue culture-infective doses) of CAEV-63 at 1 through 5 days of age (6). Serum 05 is from goat 85G05, fed colostrum containing supernatant from uninfected GSM cells at 1 through 5 days of age (6). Serum 104 is a polyvalent, monospecific goat serum prepared by immunizing goat 80G104 with gp135 purified from immunodiffusion precipitin lines (1).

Nucleotide sequence accession number. Sequence data have been deposited with the EMBL/GenBank data libraries under accession number M60855 for CAEV-63 *env*.

RESULTS

Genetic structure and variability of the CAEV-63 and CAEV-Co *env* genes. Figure 2 compares the nucleotide sequences of the *env* genes of CAEV-63 and CAEV-Co from the first of three potential ATG initiation codons. These data confirm the previously published CAEV-Co *env* gene sequence (20, 51) and show that each of the three ATG codons specifies a single ORF capable of encoding the expected SU and TM *env* gene products for both CAEV variants. The first ATG was previously identified in CAEV-Co as the initiation codon of the L or *rev* ORF and was suggested as the initiation codon of *env* (51). However, there are two additional potential initiation codons at positions 22 and 61, and the ATG at position 61 is in the context of an ACC ATGG consensus sequence identified as the optimal sequence for initiation by eukaryotic ribosomes (28). Also, a vaccinia virus P7.5-promoted construct expressed an apparently full-length recombinant CAEV SU initiated from the ATG at position 61 (30).

The 5′-terminal region of the CAEV *env* gene encodes a deduced product with the properties of a signal peptide, as illustrated by the presence of a typical signal peptide hydrophobic domain (Fig. 2). The SIGNAL sequence analysis program (16) predicts a signal peptide cleavage site after nucleotide 222 of the CAEV-63 *env* ORF and after nucleotide 228 of the CAEV-Co *env* ORF. Amino-terminal amino acid sequencing of the native CAEV-63 SU was performed to confirm this unusually long signal peptide. Identification of the first 15 amino-terminal amino acids of the CAEV-63 SU matched to the CAEV-63 *env* sequence (Fig. 2) established that the signal peptide cleavage site is 240 nucleotides downstream of the first potential ATG initiation codon or 180 nucleotides downstream of the third potential initiation codon. Thus, the CAEV-63 signal peptide is either 60 or 80 amino acids long, depending on the actual initiation codon of the *env* ORF.

Additional structural features of the CAEV *env* gene identified in Fig. 2 include the SU-TM consensus cleavage site Arg-Lys-Lys-Arg (RKKR) and the termination codons within U3. A previous CAEV-Co *env* gene sequence revealed two potential RKKR SU-TM cleavage sites (51), corresponding to nucleotides 1858 to 1869 and 1885 to 1896 (SU-TM) in Fig. 2. Comparison of the CAEV-Co and CAEV-63 sequences identifies the actual cleavage site at 1885 to 1896 for both *env* genes, as shown in Fig. 2, since the upstream alternative site in CAEV-Co is eliminated in CAEV-63 by an A to G base substitution. Both CAEV-63 and CAEV-Co *env* ORFs terminate within the U3 region of the LTR. However, a frameshift deletion 17 nucleotides upstream of U3 extends the CAEV-63 *env* ORF 88 nucleotides beyond the CAEV-Co *env* termination codon.

The sequence data in Fig. 2 indicate that, following cleavage of the signal peptide, the primary translation product of the CAEV *env* gene is a structural protein complex consisting of an amino-terminal SU and a carboxy-terminal TM separated by an RKKR cleavage site. Figure 3 shows the deduced amino acid sequence of the CAEV-63 and CAEV-Co *env* gene products aligned by the FASTA sequence analysis program (43).

Identification of the signal peptide-coding domain of the CAEV-63 *env* gene (Fig. 2) permits an unambiguous determination of the deduced structure of the SU and TM of this CAEV variant. Thus, the amino acid sequence of CAEV-63 envelope proteins (Fig. 3) specifies a 550-amino-acid SU containing 21 potential N-linked glycosylation sites and a 384-amino-acid TM containing 4 potential N-linked glycosylation sites. These values, allowing 2.1 kDa per potential N-linked glycosylation site (29), correspond to masses of 107 kDa for the SU and 46 kDa for the TM. Assuming a CAEV-Co signal peptide cleavage site analogous to that of CAEV-63, the CAEV-Co SU is 549 amino acids long, owing to deletion of one amino acid corresponding to position 360 of the CAEV-63 SU. The CAEV-Co TM (355 amino acids) is 29 amino acids shorter than the CAEV-63 TM due to the extended CAEV-63 *env* gene coding domain (Fig. 2).

Several additional structural comparisons may be made from the data in Fig. 3. All 21 potential N-linked glycosylation sites of the CAEV-63 SU are conserved in the SU of CAEV-Co, and the CAEV-Co SU contains two additional sites. In addition, the CAEV-63 and CAEV-Co TM contain four conserved potential N-linked glycosylation sites. All of the 28 cysteines within the CAEV-63 and CAEV-Co SU and TM are conserved. The CAEV-63 TM contains two additional cysteines in the extended U3 coding domain. In addition to the single amino acid deletion in the CAEV-Co SU, corresponding to an arginine residue at position 360 in the CAEV-63 SU, *env* gene nucleotide sequence variability results in 62 SU amino acid changes and 34 TM amino acid changes. The CAEV SU contains four regions of high-density amino acid variability, encompassing 35 to 60% amino acid differences. These areas correspond to CAEV-63 amino acids 88 to 100, 134 to 143, 274 to 283, and 389 to 437. Based on the relationship to conserved cysteines, the SU region containing amino acids 389 to 437 may constitute a hypervariable region analogous to the V3 principal neutralization domain of the human immunodeficiency virus type 1 (HIV-1) SU (49).

Antigenic characterization of recombinant proteins expressed by the CAEV *env* gene. Immunoprecipitations were done to demonstrate that the sequenced CAEV *env* gene clones express recombinant products that were antigenically cross-reactive with native CAEV SU and TM envelope proteins. Sera from goats vaccinated with recombinant vaccinia virus rWR-63, which expresses the CAEV-63 *env* gene derived from pC63-14B cDNA (30), were used to immunoprecipitate [^{35}S]methionine-labeled antigens of CAEV-63

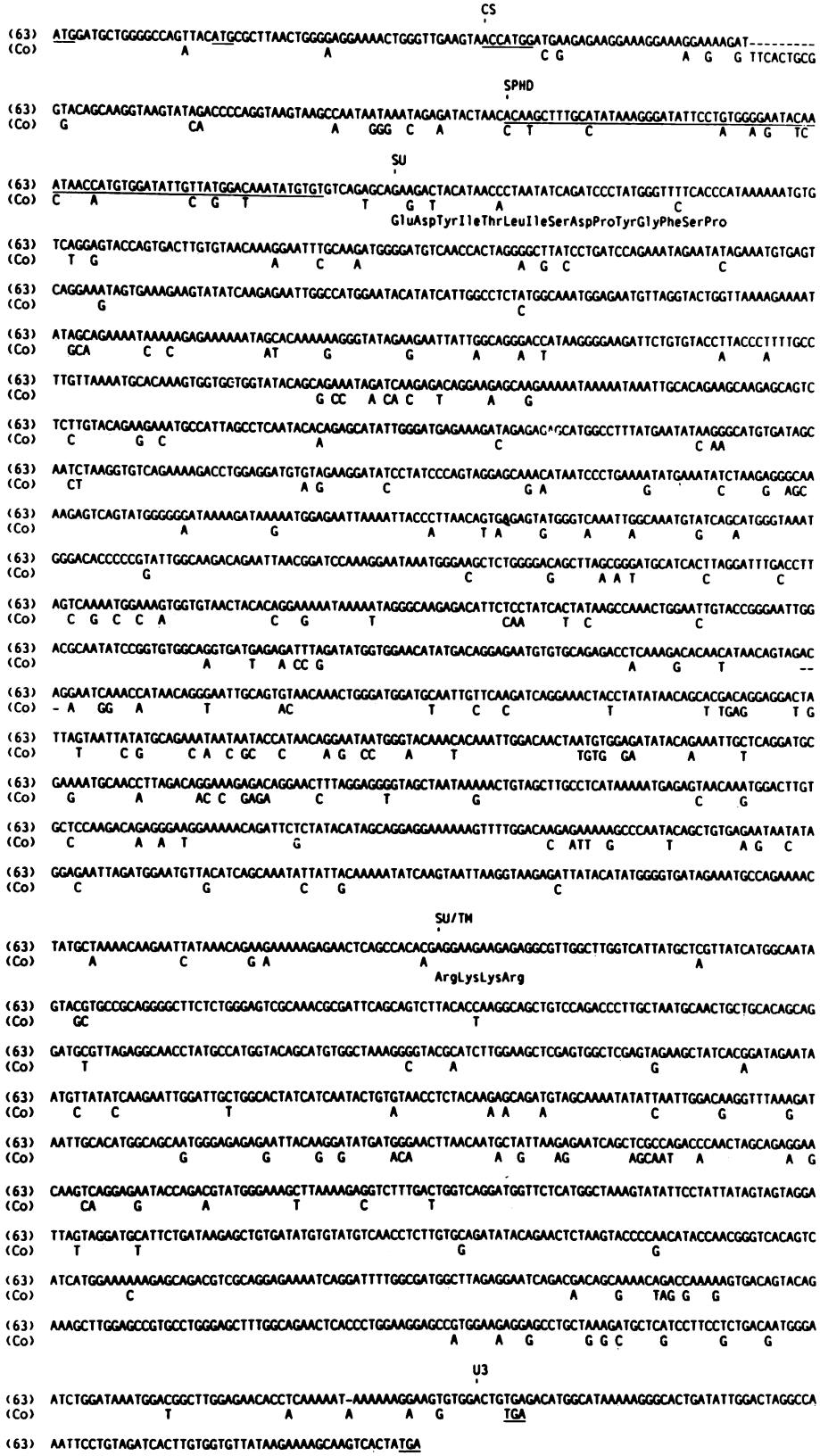

FIG. 2. Comparison of nucleotide sequences of the *env* genes of CAEV-63 and CAEV-Co from the first of the three potential ATG (underlined) initiation codons. CS, Kozak consensus sequence. SPHD, signal peptide hydrophobic domain; 29-amino-acid SPHD is underlined. SU, 5′-terminal codon of surface protein; the first 15 amino acids of SU, as determined by amino-terminal sequencing of native CAEV-63 SU, are indicated. SU/TM, surface protein-transmembrane protein Arg-Lys-Lys-Arg cleavage site. U3, initial nucleotide of the U3 region of the LTR. The TGA termination codons of the CAEV-63 and CAEV-Co *env* ORFs are indicated.

```
     SU
     |
(63) EDYITLISDPYGFSPIKNVSGVPVTCVTKEFARWGCQPLGAYPDPEIEYRNVSQEIVKEVYQEN ( 64)
(Co)                              K                 V

(63) WPWNTYHWPLWQMENVRYWLKENIAENKKRKNSTKKGIEELLAGTIRGRFCVPYPFALLKCTKW (128)
(Co)                         MQ  QQ   N  E

(63) CWYTAEIDQETGRARKIKINCTEARAVSCTEEMPLASIHRAYWDEKDRESMAFMNIRACDSNLR (192)
(Co)      A NN S K G              D      Q                 K

(63) CQKRPGGCVEGYPIPVGANIIPENMKYLRGQKSQYGGIKDKNGELKLPLTVRVWVKLANVSAWV (256)
(Co)          M         E   S      A                   L         E

(63) NGTPPYWQDRINGSKGINGKLWGQLSGMHHLGFDLSQNGKWCNYTGKIKIGQETFSYHYKPNWN (320)
(Co)       D            N T   E NS      A             E  L     Q

(63) CTGNWTQYPVWQVMRDLDMVEHMTGECVQRPQRHNITVDRNQTITGNCSVTNWDGCNCSRSGNY (384)
(Co)              I N                    G- G       T

(63) LYNSTTGGLLVIICRNNNTITGIMGTNTNWTTNVEIYRNCSGCENATLDRKETGTLGGVANKNC (448)
(Co)     SE    L L  Q S L R L         MWG  K              NTGE

(63) SLPHKNESNKWTCAPRQREGKTDSLYIAGGKKFWTREKAQYSCENNIGELDGMLHQQILLQKYQ (512)
(Co)                      D                  I   F   S   Q

                                          |TM
(63) VIKVRDYTYGVIEMPENYAKTRIINRRKRELSHTRKKRGVGLVIMLVIMAIVRAAGASLGVANA (576)
(Co)      A                       K      K                  A

(63) IQQSYTKAAVQTLANATAAQQDALEATYAMVQHVAKGVRILEARVARVEAITDRIMLYQELDCW (640)
(Co)  ___                         V

(63) HYHQYCVTSTRADVAKYINWTRFKDNCTWQQWERELQGYDGNLTMLLRESARQTQLAEEQVRRI (704)
(Co)       I   KTE                          G     T   I   K   AM        A

(63) PDVWESLKEVFDWSGWFSWLKYIPIIVVGLVGGILILRAVICVCQPLVQIYRTLSTPTYQRVTVI (768)
(Co)  E      D  ____________________________L                   E____

(63) MEKRADVAGENQDFGDGLEESDDSKTDQKVTVQKAWSRAWELWQNSPWKEPWKRSLLKMLILPL (832)
(Co)   T                       N E SER                              G  RL V

(63) TMGIWINGRLGEHLKNKKGSVDCETWHKKGTDIGLGQIPVDHLWCYKKSKSL              (884)
(Co)         W    H    RKGGL
```

FIG. 3. Comparison of deduced amino acid sequence of the *env* genes of CAEV-63 and CAEV-Co from the beginning of SU. Potential N-linked glycosylation sites are indicated by *. Cysteines are in boldface type and underlined. TM hydrophobic transmembrane domains are underlined. Regions of high-density substitutions are amino acids 88 to 100, 134 to 143, 274 to 283, and 389 to 437.

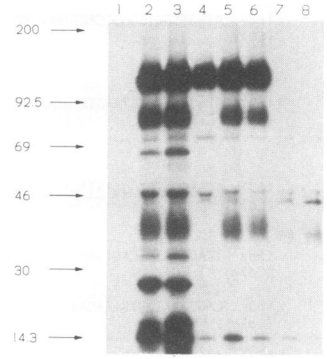

FIG. 4. Immunoprecipitation of [^{35}S]methionine-labeled CAEV with goat sera. Lane 1, serum from uninfected goat 85G05. Lanes 2 and 3, sera from CAEV-infected goats 83G79 and 85G27, respectively. Lane 4, monospecific anti-gp135 SU serum from goat 80G104. Lanes 5 and 6, sera from goats 90G65 and 90G66 immunized with rWR-63, respectively. Lanes 7 and 8, sera from goats 90G09 and 90G62 immunized with rWR-SC11, respectively. The migration of ^{14}C-labeled marker proteins is indicated (in kilodaltons).

(Fig. 4). Sera from rWR-63-vaccinated goats 90G65 and 90G66 immunoprecipitated CAEV antigens of 135, 90, and 38 kDa (lanes 5 and 6). These proteins, in addition to the 28-, 19-, and 16-kDa CAEV core proteins, were also precipitated by sera from CAEV-infected goats 83G79 and 85G27 (lanes 2 and 3). Monospecific anti-CAEV SU serum from goat 80G104 reacted only with the CAEV gp135 SU (lane 4). Control immunoprecipitations were performed with serum from a CAEV-negative goat (85G05) and sera from two goats (90G09 and 90G62) vaccinated with the control rWR-SC11 vaccinia virus without a CAEV *env* insert. As expected, none of these sera reacted with labeled CAEV antigens (lanes 1, 7, and 8). Results similar to those in Fig. 4 were also obtained by using sera from goats immunized with recombinant vaccinia virus rWR-Co, which expresses the CAEV-Co *env* gene derived from pCo-1 (30).

DISCUSSION

Analysis of *env* gene sequences from two antigenic variants of CAEV and determination of the amino-terminal amino acids of one native CAEV SU localized *env* gene variability and defined CAEV *env* gene structural domains. Nucleotide sequence heterogeneity results in four regions of high-density amino acid variability, and one region may be analogous to the V3 hypervariable domain of the HIV-1 SU (38, 49). The deduced translation products of the CAEV *env* gene consist of a 60- to 80-amino-acid signal peptide and SU and TM structural proteins. Immunization of goats with vaccinia virus expressing the *env* gene of CAEV-63 provided specific antisera for analysis of *env* gene products.

The genetic structure of the *env* genes of the bovine immunodeficiencylike virus (15), equine infectious anemia virus (50), feline immunodeficiency virus (40, 55), HIV-1 (54), and visna virus (53) has been described. In addition, the sequence of the CAEV-Co *env* gene was recently reported (51). The sequence and immunologic data described here confirm that the CAEV *env* gene, like that of other lentiviruses, encodes a primary translation product comprised of a signal peptide followed by a surface glycoprotein complex, consisting of an amino-terminal hydrophilic SU and a carboxy-terminal TM, defined by characteristic hydrophobic transmembrane domains. The RKKR SU-TM cleavage site is typical of the basic amino acid motif (RXKR, where X is any amino acid) defining other retrovirus *env* gene SU-TM cleavage sites (14).

Translation of the CAEV *env* precursor protein is initiated from one of three potential start codons. The first potential start codon in CAEV-Co has been shown to initiate *rev* (51) and was suggested as the start codon for *env* (51). However, the third potential *env* start codon is encompassed by the Kozak consensus sequence (28) and initiates an apparently full-length, biologically active recombinant CAEV SU (30). Depending on the start codon utilized by the virus, the CAEV *env* gene encodes a signal peptide of 60 to 80 amino acids. Among lentiviruses, only CAEV, visna virus, and feline immunodeficiency virus encode similar-length signal peptides (40, 53). This is in contrast to the signal peptides of equine infectious anemia virus and HIV-1, which range from 6 to 30 amino acids (2, 3, 41). The function of the additional amino acids of CAEV, visna virus, and feline immunodeficiency virus signal peptides in glycoprotein processing has not been defined.

Immunoprecipitation of native CAEV antigens with sera from goats immunized with recombinant vaccinia viruses demonstrated that antibody responses induced by the recombinant products encoded by the sequenced *env* gene clones specifically recognize native CAEV SU and TM envelope proteins. This result confirms that the CAEV *env* gene products expressed in vaccinia virus are antigenically

cross-reactive with analogous virion envelope proteins and that the CAEV SU and TM are encoded by the *env* gene. The discrepancy between the observed (135 kDa) and predicted (63 kDa) molecular masses of the SU is accounted for in part by the 21 potential N-linked glycosylation sites present in the deduced amino acid sequence. When allowances for carbohydrate are considered (2.1 kDa per site) (29), the calculated (107 kDa) and observed (135 kDa) masses of the SU agree more closely.

Whereas the CAEV SU has been well characterized (25), the properties of the TM have remained obscure. The apparent molecular masses of the TM of equine infectious anemia virus (45 kDa) and HIV-1 (41 kDa) are well established (14, 56). The TM-coding domain of the CAEV *env* gene, including potential N-linked glycosylation sites, specifies a 46-kDa gene product, and a 45-kDa protein was recently proposed as the CAEV TM (51). However, the CAEV *gag* precursor protein has a similar mobility in SDS-PAGE (9), and sera from CAEV-infected goats contain antibodies to both *gag*- and *env*-encoded proteins (17, 25). In this regard, the present data clearly show that sera from goats immunized with vaccinia virus expressing the CAEV-63 *env* gene immunoprecipitate a 90-kDa protein which is antigenically unrelated to the gp135 SU and which comigrates with a 90-kDa protein immunoprecipitated by sera from goats infected with CAEV. Sera from vaccinated and infected goats also react with a 38-kDa protein which is also unrelated to the SU. In other experiments, a monoclonal antibody against the virion 90-kDa protein immunoprecipitates only the 90- and 38-kDa proteins (35). Collectively, these experiments indicate that the 90-kDa protein is the CAEV *env*-encoded TM, which is extracted mainly as an aggregate under the conditions described here. Aggregation of the CAEV TM appears to be similar to that described for the HIV-1, HIV-2, and simian immunodeficiency virus transmembrane proteins (45, 47).

In addition to defining *env* gene coding domains and identifying *env*-encoded virion glycoproteins, the data presented localize *env* gene variability between two antigenic variants of CAEV. Genetic variation, including point mutations, deletions, and duplications of nucleotides of lentiviral *env* genes from independent isolates of bovine immunodeficiencylike virus (15), equine infectious anemia virus (42), feline immunodeficiency virus (44), HIV (38, 52), and visna virus (4) has been documented. In some cases, functional antigenic differences between isolates have been mapped to specific regions of genetic variability termed hypervariable domains (38, 49). The present data reveal four regions of high-density amino acid variability in the CAEV *env* gene. One region, corresponding to CAEV-63 SU amino acids 389 to 437, has the properties of a hypervariable region between conserved cysteines, analogous to the V3 principal neutralizing domain of the HIV-1 SU (49).

Our current studies include obtaining the sequence of the *env* genes of additional antigenic variants of CAEV and mapping the immune reactivity of SU and TM epitopes with sera from arthritic goats. These studies will provide structural analysis of virion envelope glycoprotein epitopes involved in the pathogenesis of arthritis and may lead to strategies and reagents for vaccine development.

ACKNOWLEDGMENTS

We thank S. Leib and L. Norton for technical assistance.

This work was supported by grants from the National Institutes of Health (NIAMSD AR27680) and Biomedical Research Support Program, Shared Instrument Grant (1SI0RR04759-01), the U.S. Department of Agriculture (88-34116-3735 and BARD US0937-85), the Washington Arthritis Foundation (270240), and the U.S. Department of Health and Human Services (08SORR04759A).

REFERENCES

1. **Adams, D. S., R. P. Gogolewski, A. F. Barbet, and W. P. Cheevers.** 1985. Identification of caprine arthritis-encephalitis retrovirus proteins in immunodiffusion precipitin lines. J. Gen. Virol. **66**:1139–1143.
2. **Allan, J. S., J. E. Coligan, F. Barin, M. F. McLane, J. G. Sodroski, C. A. Rosen, W. A. Haseltine, T. H. Lee, and M. Essex.** 1985. Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. Science **228**:1091–1093.
3. **Ball, J. M., V. S. V. Rao, W. G. Robey, C. J. Issel, and R. C. Montelaro.** 1988. Lentivirus antigen purification and characterization: isolation of equine infectious anemia virus gag and env proteins in one step by reverse phase HPLC and application to human immunodeficiency virus glycoproteins. J. Virol. Methods **19**:265–277.
4. **Braun, M. J., J. E. Clements, and M. A. Gonda.** 1987. The visna virus genome: evidence for a hypervariable site in the *env* gene and sequence homology among lentivirus envelope proteins. J. Virol. **61**:4046–4054.
5. **Chakrabarti, S., K. Brechling, and B. Moss.** 1985. Vaccinia virus expression vector: coexpression of β-galactosidase provides visual screening of recombinant virus plaques. Mol. Cell. Biol. **5**:3403–3409.
6. **Cheevers, W. P., D. P. Knowles, T. C. McGuire, D. R. Cunningham, D. S. Adams, and J. R. Gorham.** 1988. Chronic disease in goats orally infected with two isolates of the caprine arthritis-encephalitis lentivirus. Lab. Invest. **58**:510–517.
7. **Cheevers, W. P., D. P. Knowles, Jr., and L. K. Norton.** Neutralization-resistant antigenic variants of caprine arthritis-encephalitis lentivirus associated with progressive arthritis. J. Infect. Dis., in press.
8. **Cheevers, W. P., S. M. Roberson, P. Klevjer-Anderson, and T. B. Crawford.** 1981. Characterization of caprine arthritis-encephalitis virus: a retrovirus of goats. Arch. Virol. **67**:111–117.
9. **Cheevers, W. P., T. A. Stem, D. P. Knowles, and T. C. McGuire.** 1988. Precursor polypeptides of caprine arthritis-encephalitis lentivirus structural proteins. J. Gen. Virol. **69**:675–681.
10. **Collard, M. W., and M. D. Griswold.** 1987. Biosynthesis and molecular cloning of sulfated glycoprotein 2 secreted by rat Sertoli cells. Biochemistry **26**:3297–3303.
11. **Crawford, T. B., D. S. Adams, W. P. Cheevers, and L. C. Cork.** 1980. Chronic arthritis in goats caused by a retrovirus. Science **207**:997–999.
12. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
13. **Ellis, T. M., G. E. Wilcox, and W. F. Robinson.** 1987. Antigenic variation of caprine arthritis-encephalitis virus during persistent infection of goats. J. Gen. Virol. **68**:3145–3152.
14. **Gallaher, W. R., J. M. Ball, R. F. Garry, M. C. Griffin, and R. C. Montelaro.** 1989. A general model for the transmembrane proteins of HIV and other retrovirus. AIDS Res. Hum. Retroviruses **5**:431–440.
15. **Garvey, K. J., M. S. Oberste, J. E. Elser, M. J. Braun, and M. A. Gonda.** 1990. Nucleotide sequence and genome organization of biologically active proviruses of the bovine immunodeficiency-like virus. Virology **175**:391–409.
16. **George, D. G., W. C. Barker, and L. T. Hunt.** 1986. The protein identification resource (PIR). Nucleic Acids Res. **14**:11–15.
17. **Gogolewski, R. P., D. S. Adams, T. C. McGuire, K. L. Banks, and W. P. Cheevers.** 1985. Antigenic cross-reactivity between caprine arthritis-encephalitis, visna and progressive pneumonia viruses involves all virion-associated proteins and glycoproteins. J. Gen. Virol. **66**:1233–1240.
18. **Gubler, U., and B. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25**:263–269.
19. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p.

109–135. In D. M. Glover (ed.), DNA cloning, vol. I: a practical approach. IRL Press, Washington, D.C.
20. **Hess, J. L., J. M. Pypers, and J. E. Clements.** 1986. Nucleotide sequence and transcriptional activity of the caprine arthritis-encephalitis virus long terminal repeat. Virology **60**:385–393.
21. **Hines, S. A., T. F. McElwain, G. M. Buening, and G. H. Palmer.** 1989. Molecular characterization of Babesia bovis merozoite surface proteins bearing epitopes immunodominant in protected cattle. Mol. Biochem. Parasitol. **37**:1–10.
22. **Holmes, D., and M. Quigley.** 1981. A rapid method for plasmid DNA preparation. Focus **3**:4.
23. **Hunkapiller, M. W., E. Lujan, F. Ostrander, and L. E. Hood.** 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. **91**:227–236.
24. **Jackson, M. K., D. P. Knowles, T. A. Stem, W. G. Harwood, M. M. Robinson, and W. P. Cheevers.** 1991. Genetic structure of the pol-env regions of the caprine arthritis-encephalitis lentivirus genome. Virology **180**:389–394.
25. **Johnson, G. C., A. F. Barbet, P. Klevjer-Anderson, and T. C. McGuire.** 1983. Preferential immune response to virion surface glycoproteins by caprine arthritis-encephalitis virus-infected goats. Infect. Immun. **41**:657–665.
26. **Klevjer-Anderson, P., and W. P. Cheevers.** 1981. Characterization of the infection of caprine synovial membrane cells by the retrovirus caprine arthritis-encephalitis virus. Virology **110**:113–119.
27. **Knowles, D., Jr., W. Cheevers, T. McGuire, T. Stem, and J. Gorham.** 1990. Severity of arthritis is predicted by antibody response to gp135 in chronic infection with caprine arthritis-encephalitis virus. J. Virol. **64**:2396–2398.
28. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
29. **Lennarz, W. J.** 1975. Lipid-linked sugars in glycoprotein synthesis. Science **188**:986–991.
30. **Lichtensteiger, C. A., D. P. Knowles, Jr., T. C. McGuire, and W. P. Cheevers.** Recombinant gp135 envelope glycoproteins of caprine arthritis-encephalitis lentivirus variants inhibit homologous and heterologous variant-specific neutralizing antibodies. Virology, in press.
31. **Mackett, M., G. L. Smith, and B. Moss.** 1984. General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. J. Virol. **49**:857–864.
32. **Mackett, M., G. L. Smith, and B. Moss.** 1985. The construction and characterization of vaccinia virus recombinants expressing foreign genes, p. 191–211. *In* D. M. Glover (ed.), DNA cloning, vol. II: a practical approach. IRL Press, Washington, D.C.
33. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **McGuire, T. C., D. S. Adams, G. C. Johnson, P. Klevjer-Anderson, D. D. Barbee, and J. R. Gorham.** 1986. Acute arthritis after caprine arthritis-encephalitis virus challenge exposure of vaccinated or persistently infected goats. Am. J. Vet. Res. **47**:537–540.
35. **McGuire, T. C., D. P. Knowles, Jr., and W. P. Cheevers.** Unpublished data.
36. **McGuire, T. C., L. K. Norton, K. I. O'Rourke, and W. P. Cheevers.** 1988. Antigenic variation of neutralization-sensitive epitopes of caprine arthritis-encephalitis lentivirus during persistent arthritis. J. Virol. **62**:3488–3492.
37. **Michelson, A. M., and S. H. Orkin.** 1982. Characterization of the homopolymer tailing reaction catalyzed by terminal deoxynucleotidyl transferase: implications for the cloning of cDNA. J. Biol. Chem. **257**:14773–14782.
38. **Modrow, S., B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staal, and H. Wolf.** 1987. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. J. Virol. **61**:570–578.
39. **Narayan, O., J. E. Clements, J. D. Strandberg, L. C. Cork, and D. E. Griffin.** 1980. Biological characterization of the virus causing leukoencephalitis and arthritis in goats. J. Gen. Virol. **50**:69–79.
40. **Olmsted, R. A., V. M. Hirsch, R. H. Purcell, and P. R. Johnson.** 1989. Nucleotide sequence analysis of feline immunodeficiency virus: genome organization and relationship to other lentiviruses. Proc. Natl. Acad. Sci. USA **86**:8088–8092.
41. **Payne, S. L., F. Fang, C. Liu, B. R. Dhruva, P. Rwambo, C. J. Issel, and R. C. Montelaro.** 1987. Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes reported for sequential isolates of HIV. Virology **161**:321–331.
42. **Payne, S. L., K. Rushlow, B. R. Dhruva, C. J. Issel, and R. C. Montelaro.** 1989. Localization of conserved and variable antigenic domains of equine infectious anemia virus envelope glycoproteins using recombinant *env*-encoded protein fragments produced in *Escherichia coli*. Virology **172**:609–615.
43. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85**:2444–2448.
44. **Phillips, T. R., R. L. Talbott, C. Lamont, S. Muir, K. Lovelace, and J. H. Elder.** 1990. Comparison of two host cell range variants of feline immunodeficiency virus. J. Virol. **64**:4605–4613.
45. **Pinter, A., W. J. Honnen, S. A. Tilley, C. Bona, H. Zaghouani, M. K. Gorny, and S. Zolla-Pazner.** 1989. Oligomeric structure of gp41, the transmembrane protein of human immunodeficiency virus type 1. J. Virol. **63**:2674–2679.
46. **Pyper, J. M., J. E. Clements, M. A. Gonda, and O. Narayan.** 1986. Sequence homology between cloned caprine arthritis encephalitis virus and visna virus, two neurotropic lentiviruses. J. Virol. **58**:665–670.
47. **Rey, M., A. G. Laurent, J. McClure, B. Krust, L. Montagnier, and A. G. Hovanessian.** 1990. Transmembrane envelope glycoproteins of human immunodeficiency virus type 2 and simian immunodeficiency virus SIV-mac exist as homodimers. J. Virol. **64**:922–926.
48. **Roberson, S. M., and W. P. Cheevers.** 1984. A physical map of caprine arthritis-encephalitis provirus. Virology **134**:489–492.
49. **Rusche, J. R., K. Javaherian, C. McDanal, J. Petro, D. L. Lynn, R. Grimaila, A. Langlois, R. C. Gallo, L. O. Arthur, P. J. Fischinger, D. P. Bolognesi, S. C. Putney, and T. J. Matthews.** 1988. Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino-acid sequence of the viral envelope, gp120. Proc. Natl. Acad. Sci. USA **85**:3198–3202.
50. **Rushlow, K., K. Olsen, G. Stiegler, S. L. Payne, R. C. Montelaro, and C. J. Issel.** 1986. Lentivirus genomic organization: the complete nucleotide sequence of the *env* gene region of equine infectious anemia virus. Virology **155**:309–321.
51. **Saltarelli, M., G. Querat, D. A. M. Konings, R. Vigne, and J. E. Clements.** 1990. Nucleotide sequence and transcriptional analysis of molecular clones of CAEV which generate infectious virus. Virology **179**:347–364.
52. **Simmonds, P., P. Balfe, C. A. Ludlam, J. O. Bishop, and A. J. Leigh Brown.** 1990. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1. J. Virol. **64**:5840–5850.
53. **Sonigo, P., M. Alizon, K. Staskus, D. Klatzmann, S. Cole, O. Danos, E. Retzel, P. Tiollais, A. Haase, and S. Wain-Hobson.** 1985. Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. Cell **42**:369–382.
54. **Starcich, B. R., B. H. Hahn, G. M. Shaw, P. D. McNeely, S. Modrow, H. Wolf, E. S. Parks, W. P. Parks, S. F. Josephs, R. C. Gallo, and F. Wong-Staal.** 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. Cell **45**:637–648.
55. **Talbott, R. L., E. E. Sparger, K. M. Lovelance, W. M. Fitch, N. C. Pedersen, P. A. Luciw, and J. H. Elder.** 1989. Nucleotide sequence and genomic organization of feline immunodeficiency virus. Proc. Natl. Acad. Sci. USA **86**:5743–5747.
56. **Veronese, F. D., A. L. Devico, T. D. Copeland, S. Oroszlan, R. C. Gallo, and M. G. Sarngadharan.** 1985. Characterization of gp41 as the transmembrane protein coded by the HTLV-III env gene. Science **229**:1402–1405.