# Identification and Characterization of a Spliced C-Type Lectin-Like Gene Encoded by Rat Cytomegalovirus

SEBASTIAN VOIGT, GORDON R. SANDFORD, LIJUN DING, AND WILLIAM H. BURNS\*

BMT Program, Departments of Medicine and Microbiology & Molecular Genetics, Medical College of Wisconsin, Milwaukee, Wisconsin 53226

Received 26 June 2000/Accepted 19 October 2000

The English isolate of rat cytomegalovirus (RCMV) encodes a 20-kDa protein with a C-type lectin-like domain that is expressed in the delayed-early and late phases of the viral replication cycle. Genomic sequence analysis of the restriction fragment *Kpn*R of RCMV revealed significant homology to several C-type lectin-containing molecules implicated in natural killer (NK) and T-cell interactions, as well as genes from four poxviruses and African swine fever virus. The gene is spliced into five exons and shows a splicing pattern with exon boundaries similar to those observed in the human differentiation antigen CD69. The cap site of the gene was mapped by RNase protection, 5' rapid amplification of cDNA ends, and primer extension experiments. This analysis demonstrated that the core promoter of the RCMV lectin-like gene contains a GATA rather than a TATA box. Splicing patterns were confirmed with isolates from an infected-cell cDNA library. A unique aspect of the protein is that its translation is not initiated by the canonical methionine but rather by alanine. To study its role in virus replication and pathogenesis, a recombinant virus was constructed in which the gene is interrupted. Replication in tissue culture was similar to that of wild-type virus.

Cytomegaloviruses belong to the betaherpesvirus family, whose members exhibit characteristic species specificity and salivary gland tropism (33). They generally cause a benign infection and dwell permanently in the immunocompetent host, with intermittent reactivation from latency. Two rat cytomegalovirus (RCMV) isolates are currently under investigationthe Dutch (Maastricht) isolate (9) and our English isolate (10, 35). The genomes differ in size and are genetically distinct, classifying them as different herpesviruses rather than strains (4). While sequencing the left end of the English isolate of the RCMV genome, we identified a gene starting within the first 1,000 bp from the left terminus that is homologous to C-type lectins. Such a homologue has not been previously reported in herpesviruses but has been found in poxviruses, including fowlpox (1), cowpox (39), vaccinia virus (42), and myxoma virus (11), as well as the iridovirus African swine fever virus (ASFV) (34, 43). However, the function of these viral genes is unknown.

C-type ( $Ca^{2+}$  dependent) lectins represent a group of glycoproteins which mediate cell-cell interactions as well as pathogen recognition in a  $Ca^{2+}$ -dependent fashion through their carbohydrate recognition domains (CRDs) (41). There are genes in which the CRDs are encoded by a single exon, whereas three separate exons comprise the CRDs in other genes (6). The CRD consists of 110 to 130 residues, including four cysteines that are conserved and form two disulfide bonds. Several natural killer (NK) cell receptors are known to have sequence similarity with C-type lectins, and Weis et al. designated these NK cell receptors C-type lectin-like domains (CTLDs) (41). CTLDs play a role in triggering or inhibiting NK cell-mediated cytotoxicity (for a review, see reference 27), and thus the viral homologue may be involved in immune evasion.

A striking difference between the lectin-like genes in poxviruses and the one found in RCMV that we report here is that the latter gene is spliced into five exons. Most viral homologues of host genes are unspliced and thought to represent acquisition of the spliced transcripts of these genes. Exceptions are the recently described interleukin-10 (IL-10) homologues of human CMV (HCMV) (25) and rhesus CMV (28), the G protein-coupled receptor homologues UL33 in HCMV (14) and U12 in human herpesvirus 6 (HHV6) and HHV7 (18, 32), as well as the beta chemokine homologue murine chemokine-2 (MCK-2) in mouse CMV (MCMV) (30, 36). Moreover, the lectin-like gene in RCMV is homologous to several host C-type lectin-like genes (e.g., human CD69) and presents similar exon boundaries. We investigated its role in viral replication by constructing a recombinant virus with an interrupted lectinlike gene (RCMV<sup>lec-</sup>). It is shown that wild-type RCMV (RCMV<sup>wt</sup>) and RCMV<sup>lec-</sup> grow similarly in tissue culture.

#### MATERIALS AND METHODS

Virus and cell culture. RCMV obtained from J. Hamilton (Duke University) was propagated by passage on a rat embryo fibroblast (REF) cell line. The JON isolate of RCMV was isolated by C. Smith (University of Utah). REF cells were maintained in minimal essential medium (MEM). 293T cells used for transient transfections were passaged in Dulbecco's modified Eagle's medium. All media were supplemented with 10% fetal bovine serum, gentamicin (30  $\mu$ g/ml), and 5 mM glutamine.

DNA transfections and construction of recombinant virus. A lectin-disrupted RCMV (RCMV<sup>lec-</sup>) transfer vector was constructed by inserting the *Kpn*R DNA fragment of RCMV into the *Kpn*I site of pSK Bluescript (Stratagene, La Jolla, Calif.) after the *Hin*dIII site in the multiple cloning region was mutated. The lectin-like gene was disrupted by cloning a Lox– $\beta$ -Gal–Lox expression cassette into the *Hin*dIII site within *Kpn*R. Transfections were performed using the calcium phosphate coprecipitation method (19). Subconfluent REF monolayers in TC6 plate wells were cotransfected with 2 µg of transfer vector DNA and 2 µg of virion DNA. Cytopathic effect was seen at about 5 days after transfection. Supernatants from wells showing cytopathic effect were frozen and monitored for  $\beta$ -galactosidase ( $\beta$ -Gal)-positive virus. Virus was counted on REF cells with an overlay of 0.9% methylcellulose in MEM supplemented with 5% fetal bovine serum, 25 mM HEPES, 30 µg of gentamicin per ml, and 5 mM glutamine. After

<sup>\*</sup> Corresponding author. Mailing address: MFRC 6033, 8701 Watertown Plank Rd., Milwaukee, WI 53226. Phone: (414) 456-4989. Fax: (414) 456-6533. E-mail: wburns@mcw.edu.



- $\begin{array}{c} \text{601} \quad \text{CATCCTTTCT} \underline{\text{GC}} \text{CAACGATCGATAGATCGCCCGATTGTTCATAACTATG} \\ \underline{\text{I} \quad \text{L} \quad \text{S}} \quad \overline{\text{A}} \quad \text{Q} \quad \text{R} \quad \text{S} \quad \text{I} \quad \text{D} \quad \text{P} \quad \text{P} \quad \text{I} \quad \text{V} \quad \text{H} \quad \text{N} \quad \text{Y} \\ \end{array}$
- 651 CCATTTGTCCAAAAGATTGGATCGGCCTCACAGATACGTGTTACTATTTC A I C P K D W I G L T D T C Y Y F
- 701 TCCAATTCTACAACTAACTGGACGTTCGCCCAAACCCTATGCAAAGGAAA S N S T T N W T F A Q T L C K G N
- 751 CAATTCCAACCTGGCTCACTTCAATACCGAAGAACAATACCAATTCTTAA NS N L A H F N T E E Q Y N F L
- 801 GCAGATACAAAGGAAATTTCGATTACTGGATCGGCATACACAGAGAATCA SR Y K G N F D Y W I G I H R E S
- 851 TCAGAACATCCGTGGAAATGGGCGGATAATACCACATACAACTACTCGCT S E H P W K W A D N T T Y N Y S L
- 901 TTCTATCCGGGGTGTAGAGAAATACGCCTACCTAAACGACCTCGGAATCA S I R G V E K Y A Y L N D L G I
- 951 GTAGCGCCAGGGTTTACGCGGACAAAAGGTGGTCCTGCAGCAAGTCTACA S S A R V Y A D K R W S C S K S T
- 1001 GATTCCCTCCGTTGCCAGCTTTGTTCTACATAGCTCCCTTTTCAAATATG D S L R C Q L C S T
- 1051 ATCGTCGAAGAAACTGCCCAATGTTTGTTGACGTATCCCGGTGCTAATTT
- 1101 TTTACTATGGGTAGGGTCCTTTTCAATCCTGCTAACTACACGTATCATGA
- 1151 TTGAAAAGGACCCCGCATATAGTAAAAACATACTGCGTGAGATGATGAAG
- 1201 TGTGACCGTGTCAATAAAA

a 5-day incubation, the methylcellulose was removed and the cells were washed twice with phosphate-buffered saline (PBS) and fixed in 2% paraformaldehyde–0.2% glutaraldehyde–50 mM sodium phosphate (pH 7.3) for 10 min at room temperature. Plates were then rinsed twice with PBS and stained with 100 mM sodium phosphate (pH 7.3)–1.3 mM magnesium chloride–3 mM potassium ferricyanide–3 mM potassium ferrocyanide–1 mg of X-Gal (5-bromo-4-chloro-3-

indolyl- $\beta$ -D-galactopyranoside) for 30 min at 37°C. Recombinant virus was isolated from supernatants from  $\beta$ -Gal-positive cultures by multiple cycles of limiting-dilution cultures in 96-well plates with REF cells.

tion is underlined. The translation start site in transfected 293T cells is

at positions 494 to 496 within the ORF of the protein. There is no

methionine in the entire ORF. Exon junctions are indicated by the

final base of the ending and the first nucleotide of the adjacent exon

and are double underlined. The HindIII site at position 539 depicted in

bold letters was used to generate recombinant virus by disrupting the

gene with the insertion of  $\beta$ -Gal.

**Construction of a cDNA library.** To detect possible splice sites and map the 5' end of the open reading frame (ORF), a cDNA library from REF cells infected with RCMV for 24 h was constructed using the lambda phage gt10 and Packagene systems (Promega, Madison, Wis.) according to the manufacturer's instructions. Screening of clones was performed with nick-translated (Gibco-BRL, Gaithersburg, Md.) restriction fragments of KpnR, and hybridization was performed as described earlier (10). Positive plaques were identified by double membrane lifting of the plates and isolated with the Lambda Sorb phage absorbent (Promega). The gt10 clones with the DNA of interest were subcloned into the Topo TA 2.1 vector (Invitrogen, Carlsbad, Calif.).

**RACE and DNA sequencing.** For 3' and 5' rapid amplification of cDNA ends (RACE) and DNA sequencing, polyadenylated RNA was extracted from REF cells that had been infected with RCMV for 24 h using the Quick Prep Micro mRNA purification kit (Amersham Pharmacia Biotech, Piscataway, N.J.). For 5' RACE experiments, a first gene-specific primer (GSP I), 3'-GTCGTTCGGGC CTACCGCTGA-5', located at bp 459 (Fig. 1B), was annealed to mRNA and reverse transcribed with SuperScript II reverse transcriptase (RT) (Gibco). The cDNA was purified, tailed with dCTP, and subjected to PCR with a second specific primer (GSP II), 3'-GTGTGAACTGAAACTGAAACGGAGATG-3', located at bp 273, and the abridged anchor primer 5'-GGCCACGCGTCGACTAGTACG

GGIIGGGIIGGGIIG-3' (Gibco). To map the 3' end, cDNA was amplified with the universal amplification primer and a gene-specific 5'-GTGTAGAGAAATA CGCCT-3' primer (at bp 912) (Fig. 1B). Products were amplified for 35 cycles on brane. The N-term

CGCCT-3' primer (at bp 912) (Fig. 1B). Products were amplified for 35 cycles on a Perkin Elmer 2400 thermocycler (Perkin Elmer, Foster City, Calif.), visualized by electrophoresis on a 2% agarose gel, and subsequently cloned into the vector Topo TA 2.1. The 3' and 5' cDNA ends were determined by sequencing the inserts with forward or reverse primers provided by the manufacturer on an ABI Prism 310 analyzer (Perkin Elmer).

Southern and Northern blotting. For Southern blot analysis, wild-type and recombinant viruses were digested with restriction enzyme *KpnI* or *Hind*III and run on a 0.8% polyacrylamide gel for 48 h. After overnight transfer of the DNA to Biotrans nylon membranes (ICN Biomedicals, Irvine, Calif.), the membranes were probed with an  $[\alpha^{-32}P]$ ATP-labeled nick-translated specific fragment, incubated at 67°C overnight, and washed three times before being analyzed on an Instant Imager (Packard, Meriden, Conn.).

Total RNA was harvested from RCMV-infected and uninfected REF cells with Trizol isolation solvent (Gibco). For infection, a multiplicity of infection (MOI) of 1.0 PFU of RCMV per cell was used. After 1 h of adsorption, medium was replenished, and the cultures were harvested at 24 h postinfection. To monitor for immediate-early gene expression, cycloheximide (50 µg/ml) was added to the medium 30 min prior to and at the time of infection of REF cells, and the cultures were harvested at 1, 3, 4, and 6 h postinfection. In order to distinguish between early and late viral genes, medium was supplemented with phosphonoformic acid (PFA; 300 µg/ml) at the time of infection, and RNA was extracted after 4, 8, 12, 16, 20, and 24 h. Total infected cellular RNA was denatured in 3 volumes of 66% formamide-8% formaldehyde-0.6× MOPS (morpholinepropanesulfonic acid) buffer (pH 7.0). Samples were heated for 15 min at 65°C and mixed with 2 volumes of cold  $20 \times$  SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate). The RNA was run on a 1% agarose gel, transferred to a nylon membrane, and probed with a 420-bp radiolabeled cDNA probe spanning exons 3 to 5 of the lectin ORF. RNA and [\alpha-32P]ATP-labeled DNA probes specific for RCMV immediate-early gene 1 (IE1), KpnR of RCMV, and β-actin were hybridized overnight at 45°C. Membranes were then subjected to three washes in a solution containing 5 mM sodium phosphate, 1 mM EDTA, and 0.2% sodium dodecyl sulfate (SDS) on an orbital shaker and finally exposed to Kodak XAR5 film. RNA from mock-infected REF cells served as a negative control.

**RNase protection and primer extension assays.** RNase protection assays were performed with total cellular infected RNA and control tRNA hybridized to riboprobes labeled with  $[\alpha^{-32}P]$ CTP at 58°C overnight. Unprotected RNA was digested with RNases A and T<sub>1</sub> (Gibco) for 30 min at 30°C. Primer extension was carried out with the 30-mer oligonucleotide 3-'-TTCCTCGTCGGCGGCGTTGG CGCGCTTCTGT-5' (bp 83 in Fig. 1B) that was end labeled with  $[\gamma^{-32}P]$ ATP and T4 polynucleotide kinase for 1 h at 37°C. Ten micrograms of total RNA extracted from REF cells after a 24-h infection period with RCMV was hybridized to 10<sup>4</sup> cpm of the labeled primer, as was RNA extracted from mock-infected REF cells. Reverse transcription was performed at 45°C for 30 min by adding 30 U of Thermoscript RT (Gibco). Protected and extended products were run on a 6% polyacrylamide–7 M urea sequencing gel along with a separate sequence reaction that was initiated by the corresponding primer to map the transcription start site. Sequencing was done with Sequenase (United States Biochemical Corp., Cleveland, Ohio) using *Kpn*R of RCMV as a DNA template.

Immunoprecipitation and Western blot analysis. The ORF of the cDNA library was restriction digested from the Topo TA vector and subcloned in frame into the EcoRI and BamHI sites of the Flag CMV 5a vector (Sigma, St. Louis, Mo.), with the Flag epitope at the carboxy end of the fusion protein. A total of 106 293T cells were transiently transfected (one flask with an irrelevant plasmid was used as a negative control), and cells were lysed with buffer containing 150 mM NaCl, 1% NP-40, 0.1% SDS, 50 mM Tris (pH 7.5), and 1 mM phenylmethylsulfonyl fluoride after 48 h. Following three freeze-thaw cycles, cell debris was removed by centrifugation, and the protein mixture in the supernatant was incubated with protein G-agarose that had been coupled to the Flag M2 monoclonal antibody (Sigma) overnight. After five washes, precipitates were resolved by SDS-polyacrylamide gel electrophoresis (PAGE) using the Protean MS 3 system (Bio-Rad, Hercules, Calif.). Proteins were applied to polyvinylidene difluoride (PVDF) membranes by semidry transfer at 15 V for 30 min. After blocking in PBS containing 0.1% Tween and 5% bovine serum albumin for 1 h and three subsequent washes, membranes were incubated with the M2 anti-Flag primary antibody before being exposed to a secondary horseradish peroxidaseconjugated anti-mouse immunoglobulin G (IgG) antibody. The membranes were developed using enhanced chemiluminescence (Amersham) according to the manufacturer's instructions.

Protein purification and N-terminal sequencing. Transfected 293T cells were lysed, and the fusion protein was enriched on a Flag M2 affinity matrix (Sigma).

Eluted fractions containing the protein were pooled, dialyzed overnight against PBS, and subjected to SDS-PAGE with subsequent blotting onto a PVDF membrane. The N-terminal amino acid of the protein was identified by Edman degradation (Medical College of Wisconsin Protein and Nucleic Acid Shared Facility).

**One-step growth curves.** REF cells were infected with RCMV<sup>lec-</sup> and RCMV<sup>wt</sup> at MOIs of 10 and 0.1. Cells infected at a high MOI were harvested after 2, 12, 24, 36, 48, and 72 h, and those infected at a low MOI were harvested after 1, 2, 3, 4, and 6 days.

Nucleotide sequence accession number. The sequence of the KpnR fragment of RCMV has been deposited in GenBank under accession number AF302184.

#### RESULTS

Identification of homology. The lectin-like gene of RCMV spans approximately 2 kb of genomic DNA within the KpnR fragment of RCMV (10) (Fig. 1A). The fragment was sequenced, and the results were aligned using DNA analysis software (DNAStar, Madison, Wis.). ORFs were identified, and a homology search was performed with the Blast program (2). The sequence of the RCMV lectin-like gene aligned significantly with those of the CTLDs of the murine C-type lectin DCL1 (score 164, E value le-39) (Gorski et al., unpublished), the human lectin-like transcript LLT1 (score, 125, E value, le-27) (7), the early activation antigens in both human and mouse CD69 (scores, 96 and 93, E values, 8e-19 and 5e-18, respectively) (20, 29, 37, 38, 46), the DAP12-associated lectin (3), the rat NKG2-C and NKR-P1 proteins (5), and the human type II integral membrane protein NKG2-A (21), all of which are NK cell receptor sequences with the exception of the murine DCL1 C-type lectin, which is expressed on bone marrow-derived dendritic cells. Similar indicative alignments were retrieved by restricting the Blast database to viruses. Four poxviruses (1, 11, 17, 39) as well as ASFV (34, 43), which all contain nonspliced CTLDs, showed sequence homology to the RCMV lectin. The highest score was obtained with ORF 239 of fowlpox virus (score, 91, E value, 3e-18), followed by gene D9L of cowpox virus (score, 86, E value, 7e-17) as well as multiple isolates of ASFV. To a lesser extent, homology was seen with the A40R protein of vaccinia virus and M121R/M122R of myxoma virus. Homologous sequences were aligned using the Clustal X program (Fig. 2). There are six conserved cysteines in the C terminus that form the three disulfide bonds characteristic of C-type lectins (41). In contrast to most host genes, the RCMV lectin lacked two conserved cysteines that form one of these bonds. Similarly, all the viruses that contain a CTLD except fowlpox virus are missing these two conserved cysteines (Fig. 2B).

5'-end mapping of the RCMV lectin. We were unable to identify a TATA box upstream of the cap site. Instead of a TATA box, we identified a GATA sequence (Fig. 1B) as the putative core promoter for the RCMV lectin. This assignment was verified by three methods. First, a cDNA library was made from 24-h-infected REF cells, and sequencing of cDNA isolates revealed that the transcript was spliced from at least five exons. The cDNA that was isolated was not full length, as indicated by accompanying 5' RACE experiments which showed the same splicing junctions as the cDNA library isolates but continued further 5' sequence. Primer extension as well as RNase protection assays (not shown) revealed a single band at the C nucleotide of the mRNA (Fig. 3). This was also the first base after the 5' RACE anchor primer. Therefore, a GATA box situated 30 bp upstream of the start site represents the putative promoter (Fig. 1).

A					
RCMV Lectin DCL1 LLT1 Mouse CD69 Human CD69	1:AHEDYIR.F. 38:VQGKCFR.I 18:ANPGCLHSKI 25:TSIHPEKHHI 25:TSPHPSTRH	ASVDSNTKLYCCYVVI ISTVSPVKLYCCYGVI EHSIKATLIWRLEFE EGSIQVSIPMAVLIV EGSFQVPVLCAVMNV	FLLTVVIITLSTIL, MVLTVAVIALSVAL; MFLTIIVCGMVAAL; /LITSLITALIALNV /FITILIIALALSV	SAQ <mark>RS</mark> STKKT SAIRANC GKYNCPGLYE GQYNCPGQYT	IDP <b>P:</b> 47 EQII: 84 HQE <b>P:</b> 67 KLES: 77 FSM <b>P:</b> 77
NKR-P1Rat	27:LPPDAC <mark>R</mark> CP	R <mark>S</mark> HRLAL <mark>KL</mark> S <b>C</b> AG <b>UTI</b>	LV <b>L</b> ALVGMSTLVRV	LVQXPSVEPCRVLI	⊋ENLSK <mark>T</mark> : 86
RCMV Lectin DCL1 LLT1 Mouse CD69	48:IVHNYAICP 85:INXTYAACS 68:SVCLQAACP 78:SPHHVATCKI	KDWIGLTDTCYYFSNS KNWTGVGNKCEYFSG ESWIGFQRKCEYFSTI NEWISYKBTCYDFST	STTNWTFAQTLCKGNI (PRNWTFAQAFCMAQI DTKNWTSSQRFCDSQI TKSWALAOBSSSED	NSNIAHFNTEEQYN 2AQIARFDNEEELI DADIAQVESFOELN ACTIAUIDSEKDMT	<b>HSRYKG:</b> 107 <b>KREKG:</b> 144 <b>LLRYKG:</b> 127
Human CD69 NKR-P1Rat	78:SDSHVSSS 87:GSPAKLKCP	EDWVGYQRKCYFIST KDWLSHRDKCFHVSQ *	/KRSWTSAQNAESEH SITWKESLADGGCK( *	GATIAVIDSEKDMN GATILLVQDQEELR	KRYAG:137 RNLTK:146
RCMV Lectin DCL1 LLT1 Mouse CD69 Human CD69 NKR-P1Rat	108:NFDYWIG 145:DFDCWIG 128:PSDHWIG 138:ELEHWIG 138:REEHWUG 147:RISSSFWIG	IHRESSEHPWKWADN HRESSEHPWKWTNN SREOG.OPWKWING KNEAN.QTWKWANGF KKEPG.HPWKWSNGF SYTL <mark>S</mark> DENWKWING	TYN.YSLSTREVEK EYN.NMNPILEVGR GENT.RQFPILEAGE (EFN.SWFNTTCSGR (EFN.NWFNTTCSDK TLNSDVLSITGDTE)	YAYLNDLGISS YAYLSSDRISS CAYLNDKGASS CVSVNHKNYTA CVELKNTEVSS KDSCASYSQDKYLS	ARVYADK:161 SRSYINA:198 ARHYTER:180 VDCEANF:190 MECEKNL:190 ESCDSDN:206
RCMV Lectin DCL1 LLT1 Mouse CD69 Human CD69 NKR-P1Rat	162:RWSCSKSTD 199:MWICSKLNN 181:KWICSKSD. 191:HWVCSKPSR 191:YWICNKPYK 207:IWVCQKELK	S. LRCOLCST: 179 YNLHCO TPPV: 217 		t	t
D	×	t t			
<b>B</b> RCMV Lectin Cowpox virus Fowlpox viru African swin Cowpox virus Vaccinia vir	: D9L as ORF 239 ae fever virus c2L cus A40	53:AICPKDWIGLT 1:46:LYCKECWGCYN 64:FYCPKDWGGYN 42:KICPKMTISIN 54:EDCPTDWISYN *	DTCYYFSNSTTNWTH KNCYFFSEEKNNKSI NVCYYFNNDSKNYT MICYYFSEDTTNWTA NKCIHLSTDRKTWEH *	AQTLCKGNNSN AVERCKDMDGH ATNSCKQLNST SFEYCKNMNGR CRNTCKALNPNSD *	AHFNTEEQYN:100 THIK.: 5 TSISSKEEFK: 93 ANNDT:106 ACI: 82 IKIETPNELS:103
RCMV Lectin Cowpox virus Fowlpox viru African swir Cowpox virus Vaccinia vir	3 D9L 18 ORF 239 19 fever virus 5 C2L 19 A40	101: FLSRYKGNFDY 6: IR.RYKGTFDY 94: FWLRYKGPGMF 106:N 82: 104: FLRSLRRG	WIGIHRESSEHP.W WVGIYRTNSHSP.W WIGIEKVDFNGT.W LINITKVYHHDK.LY INDTYQDTQ. WVGESEILNQTTPY	KMADN.TTYNYSUS LNCDN.KPYNNIP Kledg.ssydniyp MVN.YSLNDNFS NMIAKNATKNGNTF	IRGVEKYAYLN:148 IRGVEDYDVLN: 52 IKGIGDCAYLS:141 S.IRNSTYEK:142 IQGYI.: 96 AQRILPNCIR:151
RCMV Lectin Cowpox virus Fowlpox viru African swir Cowpox virus	s D9L 15 ORF 239 ne fever virus s C2L	149:DLGTSSARVY 53:M.GISTARIY 142:DRSIMSSFC 143:RSKY	AD <mark>KRWSC</mark> SKSTDSLR AD <mark>RRWICEVPMVDY.</mark> LPKKWICRIILL LPLLBICSK	CQLCST:179 75 163 155	

FIG. 2. Clustal X alignment of the RCMV lectin-like gene and selected host C-type lectin-like sequences (A) and viruses containing a CTLD (B). Sequences were aligned using a Gonnet 250 matrix. Conserved cysteines are marked with asterisks. The two cysteines which are not always conserved are marked by two thick arrows. The RCMV lectin contains two cysteines at the end of the protein which are marked by thin arrows. (A) The RCMV lectin most closely resembles the murine DCL1 gene, which was cloned from bone marrow-derived dendritic cells. Homology is most notable in the CTLDs. (B) Strongest homology can be seen between the RCMV lectin-like gene and ORF 239 of fowlpox virus. The two genes from cowpox virus only span either the beginning (C2L) or the end (D9L) of the CTLD. The isolate of ASFV shown here is Pretoriuskop/96/4. Gene A40R of vaccinia virus displays restricted homology. The accession numbers for the sequences depicted are AAD22055 (DCL1), AAF22159 (LLT1), P37217 (mouse CD69), CAA80298 (human CD69), A35917 (rat NKR-P1), AAF44583 (ORF 239 of fowlpox virus), CAA72585 (D9L of cowpox virus), CAA64081 (C2L of cowpox virus), AAC28421 (ASFV), and P21063 (protein A40 of vaccinia virus).

152:VTLYNNYTTELSYNYFG......168

\*

t t

Exon boundaries of human CD69 and the RCMV lectin are similar. We chose to compare the RCMV lectin with the human CD69 gene because of their high homology and their splice structures. Both genes contain five exons separated by four introns. The predicted protein of CD69 has a total of 199

Vaccinia virus A40

amino acids, which is comparable to the RCMV lectin, which has a size of about 20 kDa in the deglycosylated state (see Fig. 5). The CD69 genomic DNA is about 15 kb in length (37), compared to 2 kb in the RCMV homologue. Due to splicing, the mRNA products are reduced to 1.7 kb (29) and 1.3 kb,



FIG. 3. Primer extension analysis of the RCMV lectin-like gene. The transcription start site is indicated by an arrow. The mRNA sequence from 5' to 3' is shown on the right. RNA from mock-infected cells (lane m) did not hybridize to the labeled probe. The nucleotides are indicated at the top.

respectively. In both genes, exons 3 through 5 code for the CTLDs. In contrast to the RCMV lectin-like gene, introns 1 and 2 of the CD69 gene are very large (37). There are no methionines in the ORF of the RCMV lectin, and thus translation does not appear to be AUG initiated.

**RCMV lectin expression.** The expression of the viral gene coding for the lectin-like protein was determined by Northern blot analysis. We assessed the stage at which expression of the lectin-like gene occurs by isolating total cellular RNA at 4, 8,

12, 16, 20, and 24 h (with and without PFA) postinfection (Fig. 4). In the presence of cycloheximide, no transcription was observed, as opposed to RNA hybridized to an RCMV IE probe as a control at immediate-early time points (1 to 4 hours postinfection) (data not shown). Northern blot analysis revealed bands of the expected approximately 1.3-kb size (Fig. 4). Transcription was at low levels until 16 h postinfection, when transcription increased markedly in the absence of PFA. Therefore, the hybridization pattern characterizes the lectin-like gene as delayed-early and late.

A 20-kDa protein initiated by alanine is expressed in transfected 293T cells. Prediction of the putative protein revealed a type II transmembrane protein with a C-terminal lectin-like domain. To investigate the expression of the protein and determine its size, we transiently transfected 293T cells with a plasmid containing the Flag sequence fused to the 3' end of the ORF. Transfected cells were subsequently lysed and subjected to immunoprecipitation and SDS-PAGE analysis as well as Western blotting. Since the putative protein contains five possible N-glycosylation sites, we transfected one flask of 293T cells in the presence of tunicamycin to inhibit glycosylation. We detected a 20-kDa band that corresponds to approximately 180 amino acids (Fig. 5). This band was recognized by the anti-Flag M2 antibody that also revealed the glycosylated form at approximately 33 kDa. Neither band was detected in 293T cells transfected with an irrelevant plasmid serving as a negative control. In addition, 293T cells transfected with the RCMV lectin ORF-Flag plasmid stained positive when incubated with a fluorescein isothiocyanate (FITC)-conjugated Flag antibody (data not shown), whereas control-transfected cells did not.

Since the ORF does not contain a methionine, we opted for N-terminal sequencing using Edman degradation to identify the primary amino acid of the protein being expressed by transfected cells. Fifteen residues were reproducibly obtained and matched the expected amino acid sequence. At the N terminus was an alanine corresponding to a start codon of GCC at positions 494 to 496 (Fig. 1B).



FIG. 4. Expression of the RCMV lectin. Total RNA extracts were harvested from infected cells at the indicated times and transferred to a nitrocellulose membrane, with subsequent labeling with a radioactive probe. At delayed-early and late times, the RCMV lectin was expressed at 8 and 12 h but was dominant at 16, 20, and 24 h postinfection when PFA was omitted. A mock-infected control was negative. The blot was stripped and reprobed with actin, indicating equal amounts of RNA in the samples.

Lectin-like gene not essential for viral replication in vitro. To determine whether the C-type lectin-like gene is essential for replication, we infected REF cells at approximate MOIs of 10 (Fig. 6A) and 0.1 (Fig. 6B). The growth characteristics of the viruses were similar at both MOIs.

## DISCUSSION

This is the first report of a C-type lectin-like gene in herpesviruses. The importance of this gene encoded by RCMV is suggested by the finding of similar homologues in poxviruses such as fowlpox (1), cowpox (39), vaccinia (17, 42), and myxoma (11) viruses as well as in various isolates of the iridovirus ASFV (34). Many host genes encode similar lectin-like domains that are expressed as cell surface receptors on lymphoid cells and macrophages. These include the murine and human activation antigen CD69 (20, 29, 37, 38, 46), the rat NKR-P1 (16) and NKG2-C (5) proteins, the human type II integral membrane protein NKG2-A (21), and multiple murine Ly49 receptors (45).

The gene is located at the left end of the RCMV genome and spans approximately 2 kb of genomic DNA. The spliced mRNA comprises five exons and codes for a 20-kDa protein when not glycosylated in transfected 293T cells without the use of an AUG codon. To ensure that the lack of an AUG start codon did not occur as a result of a mutation in our laboratory strain, we sequenced the corresponding DNA stretch in the JON isolate that was isolated from a rat in Utah. This virus bears high similarity to our English isolate (10). No sequence differences between the viruses in this region were detected. To further underscore that the English and Maastricht isolates of RCMV are diverse, it is noteworthy that no lectin-like gene has been found in the latter (40).

The ORF that was cloned from the cDNA library and used for transfection contains 194 amino acids, with a mass of approximately 22 kDa. Notably, a GCC-coded alanine was identified as the N-terminal amino acid of the protein expressed after transfection of 293T cells. Although the AUG codon initiates the vast majority of proteins in eukaryotic cells (26), alternative start codons, including ACG, CUG, and AUC, which initiate translation with nonmethionine amino acids, have been reported (31). To our knowledge, this is the first finding that GCC can serve as a start codon. Notably, the 5' adjacent codon is CUG, which has previously been reported to function as a start codon (13). Its location suggests that translation might start there and the encoded leucine might serve a protective function prior to posttranslational removal. However, we have no evidence for this possibility, and our Nterminal degradation sequences reproducibly indicated alanine as the N-terminal residue.

Another unusual aspect of the RCMV lectin-like gene is that there is a GATA box located 30 bp before the cap site of the mRNA. Three different methods identified the same 5' transcript end, stressing the probability that the GATA sequence is the core promoter for this gene. GATA core promoters have been described for the UL97 homologue in guinea pig CMV (GenBank accession number U26058) and for the DNA polymerase gene of MCMV (15).

Unlike the lectin homologues in the irido- and poxviruses, the RCMV lectin-like gene is spliced. The presence of introns



FIG. 5. Immunoprecipitation and Western blot of 293T cells transfected with the Flag plasmid containing the ORF of the RCMV lectin. Cells were transfected in either the absence (lane 1) or presence of tunicamycin (1  $\mu$ g/ml) (lane 2) or transfected with an irrelevant plasmid (lane 3). After cell lysates were adsorbed with protein G-Sepharose coupled to the anti-Flag M2 antibody, an immunoblot was performed with the same antibody, and bands were detected with a secondary goat-anti mouse IgG peroxidase-labeled antibody. Sizes are indicated in kilodaltons. In lane 1, a protein was detected with a size of approximately 33 kDa, which was reduced to 20 kDa when glycosylation was inhibited (lane 2). The other two bands observed in all three lanes represent the heavy and light chains of mouse IgG.

is uncommon in viral homologues of cellular genes. An example of genetic material being acquired in the unspliced form are the IL-10 homologues in rhesus CMV (28) and HCMV (25), the MCMV chemokine MCK-2 (30, 36), and the UL33 (HCMV), M33 (MCMV), and U12 (HHV6 and HHV7) G protein-coupled receptor homologues (14, 18, 32).

The intron-exon organization in the RCMV lectin and the human CD69 gene prompted us to compare these two genes more closely. The RCMV lectin-like gene contains five exons separated by four introns, as is the case with the prototypic human CD69 gene. In both genes, the last three exons code for the CTLD (37). CTLD exon lengths are similar except for exon five in CD69, which has a longer 3' untranslated region (38). Because of the large size of the first two putative host gene introns, totaling >10 kb, it is possible that the RCMV lectin-like gene acquired the unspliced mRNA for exons 3 to 5, which encode the lectin domain as a single transcript, and separately acquired the transmembrane and cytoplasmic exons.

The manner by which viruses acquire cellular DNA or unprocessed RNA is unknown. A coinfecting retrovirus could provide the acquiring virus with the reverse transcriptase necessary for converting RNA to DNA, an event that has been described for Marek's disease virus (24). It is generally thought that spliced RNA intermediates are involved in the incorporation of cellular gene homologues into the viral genome because of the lack of cellular introns in these transcripts and the limitations on viral genome size. Although we chose to compare our homologue gene to CD69 because of their high homology and intron-exon structure, the host gene from which RCMV acquired the lectin homologue has not yet been identified.

The advantage to the virus of acquiring and maintaining the spliced version of the lectin homologue cannot be known until the functions of the RCMV lectin are elucidated. A possible



FIG. 6. Growth curves of RCMV<sup>wt</sup> and RCMV<sup>lec-</sup> viruses. REF cells were infected with RCMV<sup>wt</sup> and RCMV<sup>lec-</sup> virus at either a high MOI of 10 (A) or a low MOI of 0.1 (B). Medium was harvested at different time points, and viruses were counted by plaque assay on REF cells.

reason to maintain splicing may be the need for splice sites for regulatory events. For example, the RCMV lectin contains a splice site at the end of the transmembrane domain which might be useful for producing species of the homologue for secretion or other purposes.

Little is known about the functions of the lectin-like homo-

logues in poxviruses. The viral protein A40R in vaccinia virus localizes to the plasma membrane of infected cells (42). An A40R-deleted virus replicated as well as wild-type virus in vitro and showed no alteration in virulence in vivo. It has been hypothesized that the two myxoma virus C-type lectin homologues inhibit cell-mediated cytotoxicity by functioning as con-

stitutively active NK cell receptors when expressed on the cell surface (11). Recently, evidence has been presented that the host gene CD69 triggers NK-mediated cytolytic activity that can be depressed by the CD94 inhibitory receptor (8).

Many of the host genes with lectin-like domains are expressed on NK cells and inhibit or activate cytolytic pathways (27, 44). If the RCMV lectin is involved in modulating NK cell interaction with infected cells, it might be needed at an early stage of infection. RCMV C-type lectin-like transcripts can be detected as early as 4 h after infection but are expressed at more substantial levels at late times. However, other viral homologues involved in modulating the innate immune response, such as the HCMV MHC class I homologue UL18 (12) and MCMV MCK-2 (30, 36), are expressed at late times. In addition, two HCMV genes involved in downregulating MHC class I chains, US2 and US11, are expressed at early and late times (22, 23). Thus, viral genes that participate in the evasion of immune surveillance do not necessarily require expression immediately after infection.

Since we did not observe different growth characteristics between the lectin-disrupted recombinant RCMV and wildtype RCMV upon infection of REF cells, the lectin-like gene does not appear to be essential for viral replication in vitro. Studies are under way to determine the function of this gene during in vivo infection.

## ACKNOWLEDGMENTS

We thank S. Malarkannan and M. Stanley for comments on the manuscript.

S.V. was supported by the Deutsche Forschungsgemeinschaft.

#### REFERENCES

- Afonso, C. L., E. R. Tulman, Z. Lu, L. Zsak, G. F. Kutish, and D. L. Rock. 2000. The genome of fowlpox virus. J. Virol. 74:3815–3831.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.
- Bakker, A. B. H., J. Wu, J. H. Phillips, and L. L. Lanier. 2000. NK cell activation: Distinct stimulatory pathways counterbalancing inhibitory signals. Hum. Immunol. 61:18–27.
- Beisser, P. S., S. J. Kaptein, E. Beuken, C. A. Bruggeman, and C. Vink. 1998. The Maastricht strain and England strain of rat cytomegalovirus represent different betaherpesvirus species rather than strains. Virology 246:341–351.
- Berg, S. F., E. Dissen, I. H. Westgaard, and S. Fossum. 1998. Two genes in the rat homologous to human NKG2. Eur. J. Immunol. 28:444–450.
- Bezouska, K., G. V. Crichlow, J. M. Rose, M. E. Taylor, and K. Drickamer. 1991. Evolutionary conservation of intron position in a subfamily of genes encoding carbohydrate-recognition domains. J. Biol. Chem. 266:11604– 11609.
- Boles, K. S., R. Barten, P. R. Kumaresan, J. Trowsdale, and P. A. Mathew. 1999. Cloning of a new lectin-like receptor expressed on human NK cells. Immunogenetics 50:1–7.
- Borrego, F., M. J. Robertson, J. Ritz, J. Pena, and R. Solana. 1999. CD69 is a stimulatory receptor for natural killer cell and its cytotoxic effect is blocked by CD94 inhibitory receptor. Immunology 97:159–165.
- Bruggeman, C. A., H. Meijer, P. H. Dormans, W. M. Debie, G. E. Grauls, and C. P. van Boven. 1982. Isolation of a cytomegalovirus-like agent from wild rats. Arch. Virol. 73:231–241.
- Burns, W. H., G. M. Barbour, and G. R. Sandford. 1988. Molecular cloning and mapping of rat cytomegalovirus DNA. Virology 166:140–148.
- Cameron, C., S. Hota-Mitchell, L. Chen, J. Barrett, J. X. Cao, C. Macaulay, D. Willer, D. Evans, and G. McFadden. 1999. The complete DNA sequence of myxoma virus. Virology 264:298–318.
- Cosman, D., N. Fanger, L. Borges, M. Kubin, W. Chin, L. Peterson, and M. L. Hsu. 1997. A novel immunoglobulin superfamily receptor for cellular and viral MHC class I molecules. Immunity 7:273–282.
- Dasso, M. C., and R. J. Jackson. 1989. Efficient initiation of mammalian mRNA translation at a CUG codon. Nucleic Acids Res. 17:6485–6497.
- Davis-Poynter, N. J., D. M. Lynch, H. Vally, G. R. Shellam, W. D. Rawlinson, B. G. Barrell, and H. E. Farrell. 1997. Identification and characterization of a G protein-coupled receptor homolog encoded by murine cytomegalovirus. J. Virol. 71:1521–1529.

- Elliott, R., C. Clark, D. Jaquish, and D. H. Spector. 1991. Transcription analysis and sequence of the putative murine cytomegalovirus DNA polymerase gene. Virology 185:169–186.
- Giorda, R., W. A. Rudert, C. Vavassori, W. H. Chambers, J. C. Hiserodt, and M. Trucco. 1990. NKR-P1, a signal transduction molecule on natural killer cells. Science 249:1298–1300.
- Goebel, S. J., G. P. Johnson, M. E. Perkus, S. W. Davis, J. P. Winslow, and E. Paoletti. 1990. The complete DNA sequence of vaccinia virus. Virology 179:247–266.
- Gompels, U. A., J. Nicholas, G. Lawrence, M. Jones, B. J. Thomson, M. E. Martin, S. Efstathiou, M. Craxton, and H. A. Macaulay. 1995. The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. Virology 209:29–51.
- Graham, F. L., and A. J. van der Eb. 1973. A new technique for the assay of infectivity of adenovirus 5 DNA. Virology 52:456–467.
- Hamann, J., H. Fiebig, and M. Strauss. 1993. Expression cloning of the early activation antigen CD69, a type II integral membrane protein with a C-type lectin domain. J. Immunol. 150:4920–4927.
- Houchins, J. P., T. Yabe, C. McSherry, and F. H. Bach. 1991. DNA sequence analysis of NKG2, a family of related cDNA clones encoding type II integral membrane proteins on human natural killer cells. J. Exp. Med. 173:1017– 1020.
- Jones, T. R., L. A. Hanson, L. Sun, J. S. Slater, R. M. Stenberg, and A. E. Campbell. 1995. Multiple independent loci within the human cytomegalovirus unique short region down-regulate expression of major histocompatibility complex class I heavy chains. J. Virol. 69:4830–4841.
- Jones, T. R., and L. Sun. 1997. Human cytomegalovirus US2 destabilizes major histocompatibility complex class I heavy chains. J. Virol. 71:2970– 2979.
- Kost, R., D. Jones, R. Isfort, R. Witter, and H. J. Kung. 1993. Retrovirus insertion into herpesvirus: characterization of a Marek's disease virus harboring a solo LTR. Virology 192:161–169.
- Kotenko, S. V., S. Saccani, L. S. Izotova, O. V. Mirochnitchenko, and S. Pestka. 2000. Human cytomegalovirus harbors its own unique IL-10 homolog (cmvIL-10). Proc. Natl. Acad. Sci. USA 97:1695–1700.
- Kozak, M. 1992. Regulation of translation in eukaryotic systems. Annu. Rev. Cell Biol. 8:197–225.
- 27. Lanier, L. L. 1998. NK cell receptors. Annu. Rev. Immunol. 16:359-393.
- Lockridge, K. M., S. S. Zhou, R. H. Kravitz, J. L. Johnson, E. T. Sawai, E. L. Blewett, and P. A. Barry. 2000. Primate cytomegaloviruses encode and express an IL-10-like protein. Virology 268:272–280.
- Lopez-Cabrera, M., A. G. Santis, E. Fernandez-Ruiz, R. Blacher, F. Esch, P. Sanchez-Mateos, and F. Sanchez-Madrid. 1993. Molecular cloning, expression, and chromosomal localization of the human earliest lymphocyte activation antigen AIM/CD69, a new member of the C-type animal lectin superfamily of signal-transmitting receptors. J. Exp. Med. 178:537–547.
- MacDonald, M. R., M. W. Burney, S. B. Resnick, and H. W. Virgin IV. 1999. Spliced mRNA encoding the murine cytomegalovirus chemokine homolog predicts a beta chemokine of novel structure. J. Virol. 73:3682–3691.
- Malarkannan, S., T. Horng, P. P. Shih, S. Schwab, and N. Shastri. 1999. Presentation of out-of-frame peptide/MHC class I complexes by a novel translation initiation mechanism. Immunity 10:681–690.
- Megaw, A. G., D. Rapaport, B. Avidor, N. Frenkel, and A. J. Davison. 1998. The DNA sequence of the RK strain of human herpesvirus 7. Virology 244: 119–132.
- Mocarski, E. S. 1996. Cytomegaloviruses and their replication, p. 2447–2492. *In* B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), Fields virology, 3rd ed., vol. 2. Lippincott-Raven, Philadelphia, Pa.
- 34. Neilan, J. G., M. V. Borca, Z. Lu, G. F. Kutish, S. B. Kleiboeker, C. Carrillo, L. Zsak, and D. L. Rock. 1999. An African swine fever virus ORF with similarity to C-type lectins is non-essential for growth in swine macrophages in vitro and for virus virulence in domestic swine. J. Gen. Virol. 80:2693– 2697.
- Priscott, P. K., and D. A. J. Tyrrell. 1982. The isolation and partial characterization of a cytomegalovirus from the brown rat Rattus norvegicus. Arch. Virol. 73:145–160.
- Saederup, N., Y. C. Lin, D. J. Dairaghi, T. J. Schall, and E. S. Mocarski. 1999. Cytomegalovirus-encoded beta chemokine promotes monocyte-associated viremia in the host. Proc. Natl. Acad. Sci. USA 96:10881–10886.
- Santis, A. G., M. Lopez-Cabrera, J. Hamann, M. Strauss, and F. Sanchez-Madrid. 1994. Structure of the gene coding for the human early lymphocyte activation antigen CD69: a C-type lectin receptor evolutionarily related with the gene families of natural killer cell-specific receptors. Eur. J. Immunol. 24: 1692–1697.
- Santis, A. G., M. Lopez-Cabrera, F. Sanchez-Madrid, and N. Proudfoot. 1995. Expression of the early lymphocyte activation antigen CD69, a C-type lectin, is regulated by mRNA degradation associated with AU-rich sequence motifs. Eur. J. Immunol. 25:2142–2146.
- 39. Shchelkunov, S. N., P. F. Safronov, A. V. Totmenin, N. A. Petrov, O. I. Ryazankina, V. V. Gutorov, and G. J. Kotwal. 1998. The genomic sequence analysis of the left and right species-specific terminal region of a cowpox

virus strain reveals unique sequences and a cluster of intact ORFs for immunomodulatory and host range proteins. Virology 243:432-460.

- 40. Vink, C., E. Beuken, and C. A. Bruggeman. 2000. Complete DNA sequence of the rat cytomegalovirus genome. J. Virol. 74:7656-7665.
- 41. Weis, W. I., M. E. Taylor, and K. Drickamer. 1998. The C-type lectin superfamily in the immune system. Immunol. Rev. 163:19-34.
- 42. Wilcock, D., S. A. Duncan, P. Traktman, W. H. Zhang, and G. L. Smith. 1999. The vaccinia virus A4OR gene product is a nonstructural, type II membrane glycoprotein that is expressed at the cell surface. J. Gen. Virol. 80:2137-2148.
- 43. Yanez, R. J., J. M. Rodriguez, M. L. Nogal, L. Yuste, C. Enriquez, J. F.

Rodriguez, and E. Vinuela. 1995. Analysis of the complete nucleotide sequence of African swine fever virus. Virology 208:249–278.
44. Yokoyama, W. M. 1998. Natural killer cell receptors. Curr. Opin. Immunol.

- 10:298-305.
- 45. Yokoyama, W. M., and W. E. Seaman. 1993. The Ly-49 and NKR-P1 gene families encoding lectin-like receptors on natural killer cells: the NK gene complex. Annu. Rev. Immunol. 11:613-635.
- 46. Ziegler, S. F., S. D. Levin, L. Johnson, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, E. Baker, G. R. Sutherland, A. L. Feldhaus, and F. Ramsdell. 1994. The mouse CD69 gene: structure, expression, and mapping to the NK gene complex. J. Immunol. 152:1228-1236.