JOURNAL OF VIROLOGY, Nov. 1988, p. 4259–4264
0022-538X/88/114259-06$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 62, No. 11

Translocation of Rubella Virus Glycoprotein E1 into the Endoplasmic Reticulum

TOM C. HOBMAN, ROBERT SHUKIN, AND SHIRLEY GILLAM*

Department of Pathology, Faculty of Medicine, University of British Columbia, Research Centre, 950 West 28th Avenue, Vancouver, British Columbia, Canada V5Z 4H4

Received 14 March 1988/Accepted 27 July 1988

Rubella virus (RV) contains four structural proteins, C (capsid), E2a, E2b, and E1, which are derived from posttranslational processing of a single polyprotein precursor, p110. C protein is nonglycosylated and is thought to interact with RV RNA to form a nucleocapsid. E1 and E2 are membrane glycoproteins that form the spike complexes located on the virion exterior. Two different E1 cDNAs were used to analyze the requirements for translocation of E1 into the endoplasmic reticulum. Analysis of expression of these cDNAs both in vivo and in vitro showed that RV E1 was stably expressed and glycosylated in COS cells and correctly targeted into microsomes in the absence of E2 glycoprotein. The results provide experimental evidence that translocation of RV E1 glycoprotein into the endoplasmic reticulum is mediated by a signal peptide contained within the 69 carboxyl-terminal residues of E2.

Rubella virus (RV), the causative agent of German measles, is a small, enveloped RNA virus whose genome consists of a single-stranded 40S RNA of positive polarity (8). Cells infected with RV, besides containing the 40S RNA, have an RV-specific 24S subgenomic RNA species which corresponds to the 3′ one-third of the genomic RNA and encodes the viral structural proteins (20). Both RNA species are polyadenylated and have 5′-terminal caps (20). The order of translation, NH_2-C-E2-E1-COOH, has been determined by synchronized translation (18) and sequence analysis of the 24S subgenomic RNA of RV (3). The strategy for the expression of RV structural proteins was first proposed by Oker-Blom et al. (18, 20) and later confirmed by cDNA sequencing and expression (3).

Translation of the 24S subgenomic RNA produces a 110-kilodalton precursor polyprotein (p110) which is processed to yield four structural proteins, C (capsid), E2a, E2b, and E1, with respective molecular weights of 33,000, 47,000, 42,000, and 58,000 (19, 20). The capsid protein is nonglycosylated and rich in arginine residues, which are thought to participate in the binding of 40S RNA to form the nucleocapsid (3). E2a, E2b, and E1 are membrane glycoproteins found on the virion surface (19). E2a and E2b are known to result from heterogeneous processing of the same E2 apoprotein moiety (10, 20).

E1 and E2 proteins contain N-linked oligosaccharide as a consequence of their passage through the endoplasmic reticulum (ER) and Golgi apparatus. According to the signal hypothesis, translocation of secretory and membrane proteins into the ER is mediated by a hydrophobic signal peptide (2). Signal peptides for E2 and E1 tentatively identified on the basis of cDNA sequence data (for E2, see references 3 and 26; for E1, see references 5, 17, and 26) are more or less in agreement in both location and sequence. E2 and E1 both contain uninterrupted stretches of hydrophobic amino acids of 19 and 27 residues, respectively, which probably function as membrane-binding domains (3, 25), making them group 1 proteins (24).

As the processing of p110 is unclear, virtually nothing is known about the targeting of E2 and E1 to the ER membrane. We have constructed two E1 cDNAs, one of which contains the putative signal peptide for E1 and one which has only the E1-coding region preceded by an ATG. By expressing these constructs in vitro and in vivo, we demonstrate that the 69 carboxyl-terminal amino acids of E2 contain a sequence necessary for the insertion of E1 into the ER.

MATERIALS AND METHODS

General recombinant DNA techniques. Restriction endonucleases and DNA-modifying enzymes were purchased from commercial sources and used according to the manufacturer's specifications. Growth and purification of recombinant plasmids were essentially as described previously (15). Plasmids were propagated in *Escherichia coli* DH5α (Bethesda Research Laboratories, Inc.). [α-^{32}P]dATP and L-[^{35}S]methionine were purchased from Amersham Corp. Plasmids pSPT19 and pSVL were purchased from Pharmacia, Inc. The oligonucleotide used for mutagenesis was obtained from the laboratory of M. Smith (University of British Columbia).

Oligonucleotide-directed mutagenesis. Site-specific mutagenesis based on the method of Kunkel (13) was used to introduce an in-frame ATG in front of the E1-coding sequence found in plasmids pE1 and pSVL-E1. Briefly, the 1,550-nucleotide (nt) *Pst*I fragment from pSTP19(C/E2/E1) (3) was subcloned into M13mp19 and used to transform the host strain, *dut ung E. coli* CJ236 (13). Uracil-containing single-stranded DNA was isolated and mutagenized with the oligonucleotide 5′GCCTATGGCATGGAGGAGGCT3′. The presence of the ATG in front of E1 was confirmed by dideoxynucleotide sequencing (21).

Recombinant plasmids. Plasmid p3′E2E1 was constructed by removal of the 1,500-nt *Nco*I fragment from pSTP19(C/E2/E1) (3) and religation, creating a cDNA encoding the capsid translation start site with eight residues from C, 69 residues from the C-terminal E2, and all of E1, including the putative signal peptide (see Fig. 1A and B). To create pE1, we excised the 1,550-nt *Pst*I fragment from pSTP19(C/E2/E1), subcloned it into M13mp19, mutagenized it as described above and subcloned the insert back into pSTP19 downstream from the phage SP6 promoter (see Fig. 1A and B). Simian virus 40-based vector plasmids (pSVL-3′E2E1, pSVL-E1) were made by excising the 3′E2E1 or E1 insert

* Corresponding author.

from the respective pSPT19 vectors with *Eco*RI and *Hin*dIII. The cohesive ends were filled in with Klenow enzyme, and the cDNA inserts were ligated into the *Sma*I site of pSVL. Constructs were then screened for orientation by restriction analysis.

In vitro transcription. pE1 or p3'E2E1 (5 μg) was linearized with *Hin*dIII, extracted with phenol-chloroform, and precipitated with ethanol. The DNA template was then added to a transcription-capping reaction mixture containing 40 mM Tris chloride (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM dithiothreitol, 500 U of RNasin (Promega Biotec) per ml, 100 μg of nuclease-free bovine serum albumin per ml, 0.5 mM each ATP, CTP, UTP, and cap analog GpppG, 0.1 mM GTP, and 600 U of SP6 RNA polymerase (Promega) per ml. Reactions were typically carried out in 50-μl volumes, and reaction mixtures were incubated for 1 h at 37°C, after which the DNA was removed by DNase I digestion. The reactions were successively extracted with phenol-chloroform and ether, precipitated with ethanol, taken up in 30 μl of water, and stored at −70°C.

In vitro translation. Translation of SP6-derived transcripts was performed in a nuclease-treated rabbit reticulocyte lysate system (Promega) containing 0.02 mM amino acid mixture minus methionine, [^{35}S]methionine at 1,200 μCi/ml, RNasin at 1,600 U/ml, and RNA at 40 μg/ml in a 25-μl volume. Some reactions were supplemented with one equivalent of canine pancreatic microsomes (Promega). After incubation at 30°C for 1 h, translation products were assayed for protection from exogenously added protease, separated by electrophoresis on 10% Laemmli gels (14), and analyzed by fluorography at −70°C.

Protease protection assay. Protease digestion conditions were essentially those described previously (28). After in vitro translation, $CaCl_2$ was added to 1 mM, and the reaction mixtures were kept on ice for 10 min to stabilize the membranes. Samples were then incubated for 60 min on ice in the presence of trypsin (300 μg/ml). Proteolysis was stopped by the addition of phenylmethylsulfonyl fluoride to 1 mM and trypsin inhibitor to 500 μg/ml and incubation on ice for 10 min. Triton X-100 was included at a final concentration of 1% in some reactions. Samples were separated on 10% Laemmli gels (14) and analyzed by fluorography.

Transfection of COS cells and metabolic labeling. COS cells were transfected with plasmids pSVL-E1 and pSVL-3'E2E1 by the method of Adams and Rose (1). Briefly, subconfluent monolayers grown in Dulbecco modified Eagle medium (DMEM) plus 5% fetal calf serum were washed twice with Tris-saline (25 mM Tris chloride [pH 7.4], 140 mM NaCl, 3 mM KCl, 1 mM $CaCl_2$, 0.5 mM $MgCl_2$, 0.9 mM Na_2HPO_4). Cells were then exposed to DEAE-dextran (M_r, 5×10^5; 500 μg/ml) and plasmid DNA (4 μg/ml) in Tris-saline at 37°C for 30 min, after which the DNA solution was removed and replaced with DMEM plus 40 μM chloroquine; the mixture was then incubated at 37°C for 3 h. Finally, monolayers were washed three times with Tris-saline and incubated at 37°C for 40 h in DMEM plus 5% fetal calf serum.

To label cellular proteins, we washed monolayers twice with DMEM without methionine but with 5% dialyzed calf serum, starved them in this medium for 30 min at 37°C, and incubated them in DMEM without methionine but with 5% dialyzed fetal calf serum and 400 μCi of [^{35}S]methionine for 3 h at 37°C. Cells were washed three times with cold Tris-saline and lysed in 500 μl of 50 mM Tris chloride (pH 7.5)–150 mM NaCl–1% Nonidet P-40. Lysates were cleared of nuclei and debris by 5 min of centrifugation in a Microfuge at 4°C.

Immunoprecipitation. Rabbit polyclonal antirubella serum (10 μl) was incubated with 100 μl of COS cell lysate plus 400 μl of NET (400 mM NaCl, 50 mM Tris chloride [pH 8.0], 5 mM EDTA)–1% Nonidet P-40 for at least 4 h at 4°C on a rotator. Immune complexes were precipitated with protein A-Sepharose for 1 h at 4°C, washed five times with NET–1% Nonidet P-40, and washed once with water. The immunoprecipitates were disrupted in 62 mM Tris chloride (pH 6.8)–10% glycerol–2% sodium dodecyl sulfate–2% 2-mercaptoethanol at 100°C for 3 min, and the supernatant was analyzed by electrophoresis through a 10% Laemmli gel (14) and fluorography.

Endo H digestion. Immune complexes still attached to protein A-Sepharose beads were suspended in 75 μl of 0.13 M sodium citrate (pH 5.5)–1 mM phenylmethylsulfonyl fluoride and incubated with or without 2 mU of endo-β-*N*-acetylglucosaminidase H (endo H) at 37°C for 14 h. The immune complexes were washed once with water, disrupted, electrophoresed, and fluorographed as described above.

RNA isolation and Northern blot analysis. Total cellular RNA was isolated from transfected COS cells as described previously (15). RNA (5 μg) was electrophoresed through a 1% agarose gel containing 6% formaldehyde and then transferred to nitrocellulose. RV E1-specific RNA was detected by hybridization to a ^{32}P-labeled 30-mer (5'GTCCTCGCCA TTGACGGTAAGATGGCAGTT3') complementary to E1.

S1 nuclease mapping. S1 nuclease protection analysis was performed by the method of Favaloro et al. (4) with modifications. ^{32}P-labeled antisense RNA from *Eco*RI-linearized p3'E2E1 was synthesized by using T7 RNA polymerase (Promega). The probe (10^6 cpm) was coprecipitated with 4 μg of total cellular RNA isolated from transfected COS cells, suspended in 20 μl of hybridization buffer (40 mM piperazine-*N*,*N*'-bis(2-ethanesulfonic acid) [PIPES] [pH 6.8], 400 mM NaCl, 1 mM EDTA, 80% formamide), heated at 80°C for 15 min, and hybridized at 65°C. After 3 h, 300 μl of ice-cold S1 buffer (280 mM NaCl, 30 mM sodium acetate [pH 4.4], 4.5 mM $ZnCl_2$, 20 μg of sonicated salmon sperm DNA per ml) containing 300 U of S1 nuclease (New England Nuclear Corp.) was added to each tube and incubated at 37°C for 30 min. Digestions were terminated by adding 75 μl of 2.5 mM ammonium acetate–50 mM EDTA–20 μg of tRNA and were precipitated overnight at −20°C after the addition of 400 μl of isopropanol. Samples were centrifuged in a Microfuge, and the pellets were dissolved in 5 μl of H_2O at 60°C and mixed with 15 μl of denaturing mix (20 mM morpholinepropanesulfonic acid [MOPS] [pH 7.0], 2 M formaldehyde, 50% formamide) heated at 80°C for 15 min and quickly cooled on ice. Samples were separated on an 0.8% agarose gel containing 2.2 M formaldehyde, transferred to Hybond N (Amersham), and autoradiographed.

RESULTS

Translocation into microsomes. For the translocation assays, two E1 cDNAs cloned downstream from the SP6 promoter in pSPT19 were utilized. Plasmid pE1 contains the entire E1-coding sequence (481 residues) and a methionine engineered directly in front of the gene (Fig. 1A and B). The construct p3'E2E1, in addition to containing the entire E1 gene, also contains the capsid translation start site as well as nucleotides specifying the first eight amino acids of C and 69 carboxyl-terminal residues of E2, including the putative E1

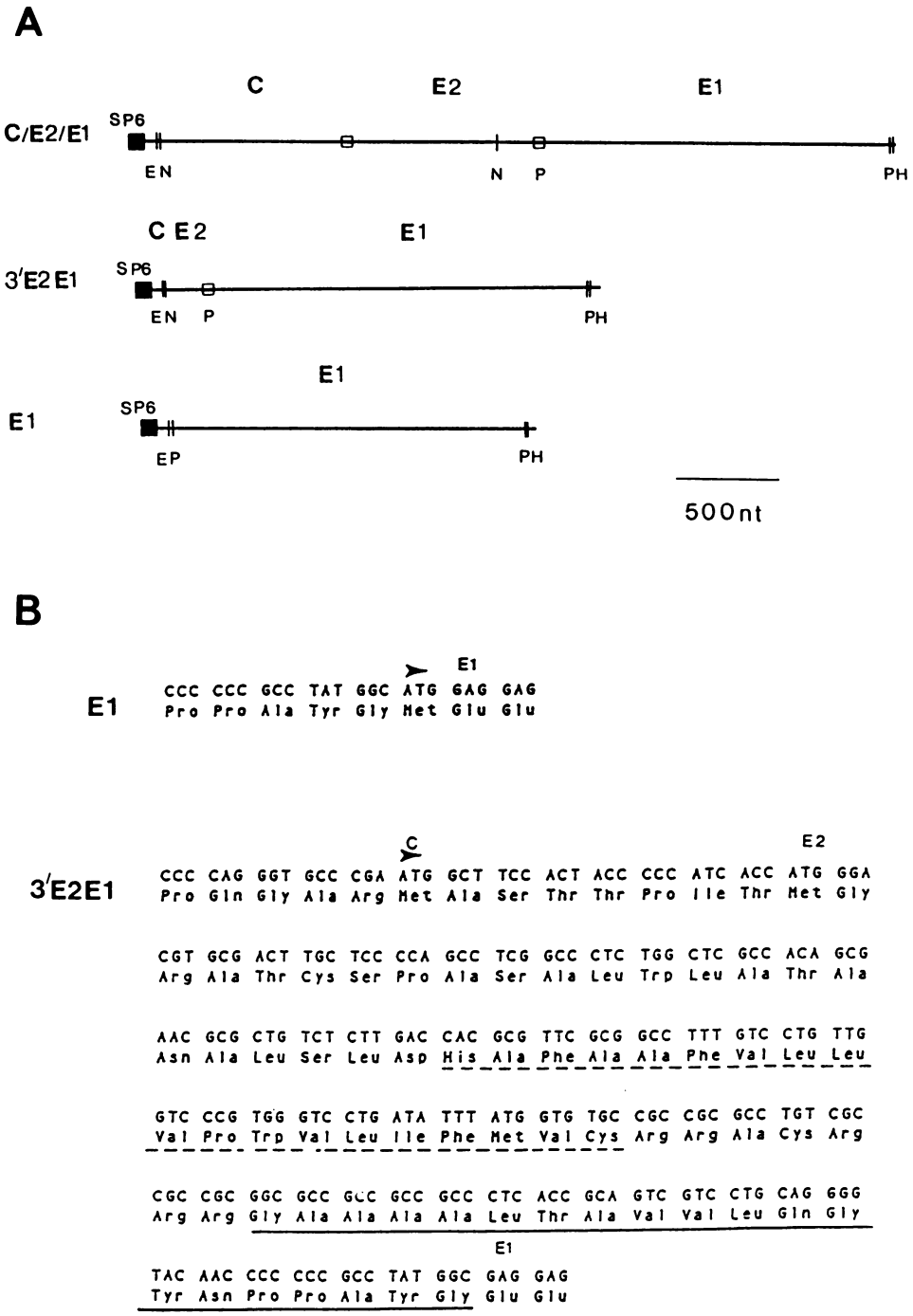

FIG. 1. (A) Physical maps of RV cDNAs. Insert C/E2/E1 codes for the structural genes C, E2, and E1 and was inserted between the *Eco*RI and *Hin*dIII sites of pSPT19 to create pSPT19(C/E2/E1). 3'E2E1 cDNA was made by deleting the 1,500-nt *Nco*I fragment from pSPT19(C/E2/E1) and is contained in plasmids p3'E2E1 and pSVL-3'E2E1. Constructs pE1 and pSVL-E1 contain the 1,550-nt *Pst*I fragment from 3'E2E1 and an in-frame ATG engineered directly in front of the E1-coding sequence. SP6 RNA polymerase promoters are marked with solid squares, and putative signal peptides for E2 and E1 are indicated by open squares. Restriction enzyme sites are shown as follows: E, *Eco*RI; N, *Nco*I; P, *Pst*I; and H, *Hin*dIII. (B) sequence of 5' regions of E1 and 3'E2E1 inserts up to the E1-coding region. Translation initiation sites are designated by arrowheads. The putative transmembrane anchor of E2 and the signal peptide of E1 are underlined with broken and solid lines, respectively. The N-terminal coding region (GAGGAG) of E1 in each cDNA is indicated. The exact carboxyl-terminal terminus of E2 has not been determined.

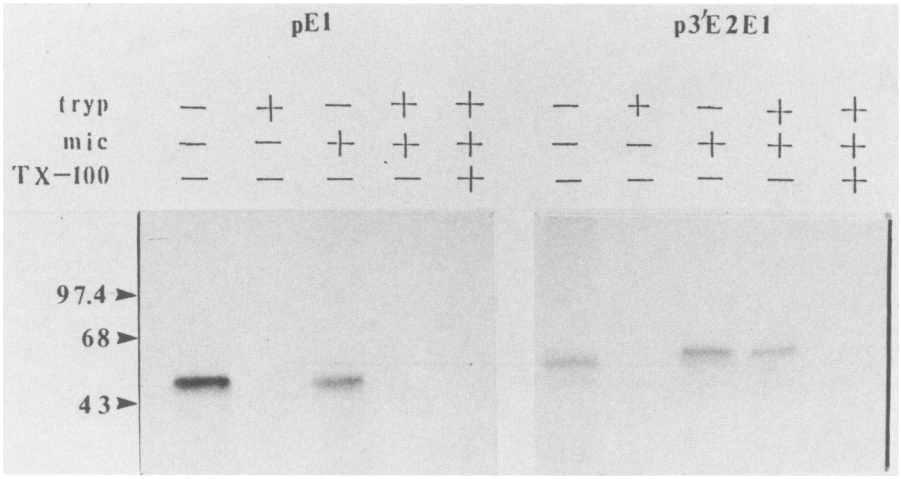

FIG. 2. Cell-free translation of pE1 and p3'E2E1 RNAs. Translation was performed at 30°C in the presence of L-[^{35}S]methionine, with or without canine microsomes (mic). Trypsin (trp) (300 μg/ml) digestion was done at 0°C for 1 h, and the products were separated on a 10% Laemmli gel and fluorographed at −70°C. Some trypsin digestions included 1% Triton X-100 (TX-100). Numbers at left are molecular weights in thousands.

signal sequence (Fig. 1A and B). Plasmids were linearized with *Hin*dIII, and capped full-length transcripts were transcribed with SP6 RNA polymerase and translated in vitro, using a nuclease-treated rabbit reticulocyte system.

In the presence of microsomes, a functional signal peptide should confer translocational competency to E1 such that the protein is protected from exogenously added protease. The translational product of pE1 is a protein which has a molecular weight of approximately 51,600 and which, when synthesized with canine microsomes present, was apparently not translocated across the membranes, as evidenced by no detectable change in molecular weight (Fig. 2). Incubation of the protein with trypsin, with or without membranes, resulted in complete proteolysis, indicating that the pE1 protein was not sequestered in the microsomes.

Cell-free translation of RNA from p3'E2E1 produced a protein with an apparent molecular weight of 59,000 to 60,000. Translation of this protein with microsomes gave rise to a slightly larger polypeptide species (Fig. 2). Presumably this increase in size was due to core glycosylation at one or more of the three potential N-linked glycosylation sites in E1 (3). Complete degradation occurred after trypsin treatment in the absence of microsomes; however, protection from proteolysis was observed after translation with microsomal membranes. No significant decrease in size was noticeable after proteolysis, indicating that most of the p3'E2E1 protein was sequestered in the vesicles. The microsome-mediated protection was abolished when trypsin digestion was performed with Triton X-100.

Expression in COS cells. The cDNA inserts from pE1 and p3'E2E1 were subcloned into the simian virus 40-based eukaryotic expression vector pSVL. The vector contains the simian virus 40 late promoter and polyadenylation site and is used primarily for high-level transient expression in COS cells. COS cells were transfected with pSVL-E1 and pSVL-3'E2E1, and the cellular proteins were labeled with [^{35}S]methionine. Fluorographs from COS cell radioimmunoprecipitates are shown in Fig. 3A. Cells containing pSVL-3'E2E1 expressed a protein that was antigenically similar to native RV E1 glycoprotein (Fig. 3A, lane 2). Immunoprecipitates from COS cells expressing pSVL-E1 contained no E1-specific protein, even after extended exposure times (Fig. 3A, lane 3). Figure 3B shows that the E1-specific protein from pSVL-3'E2E1-transfected cells was endo H sensitive and therefore contained N-linked glycans of the high-mannose variety.

Northern blot analysis. Total cellular RNA from COS cells transfected with pSVL-E1 and pSVL-3'E2E1 was prepared and electrophoresed through a denaturing agarose gel, transferred to nitrocellulose, and probed with a ^{32}P-labeled oligonucleotide complementary to E1. Both pSVL-E1- and

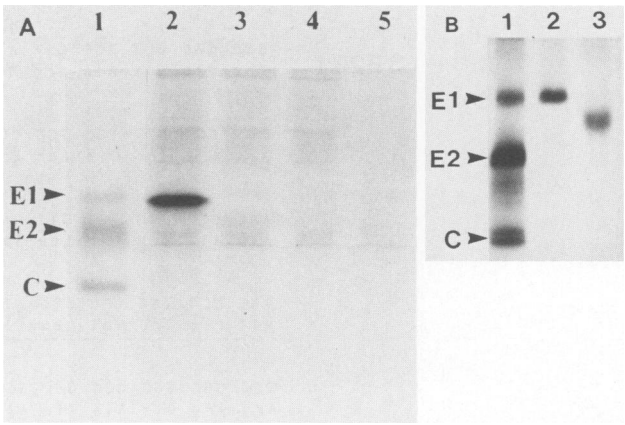

FIG. 3. (A) Expression of pSVL-E1 and pSVL-3'E2E1 in COS cells. COS cells were transfected with the respective plasmids in the presence of DEAE-dextran and grown at 37°C for 40 h. Cellular proteins were labeled with 400 μCi of L-[^{35}S]methionine for 3 h, immunoprecipitated, separated on a 10% Laemmli gel, and fluorographed. Lanes: 1, ^{35}S-labeled RV strain M33-, 2 and 5, pSVL-3'E2E1-transfected COS cell lysates; 3, pSVL-E1 COS cell lysate; 4, pSVL (vector)-transfected COS cells. Lanes 1 to 4 are radioimmunoprecipitates produced with rabbit polyclonal serum to RV strain M33; in lane 5 preimmune rabbit serum was used. (B) Deglycosylation of radioimmunoprecipitated proteins from pSVL-3'E2E1-transfected COS cells. COS cells were transfected as described in panel A. Lanes: 1, RV structural proteins; 2 and 3, immunoprecipitates from pSVL-3'E2E1-transfected COS cells. In lane 3, the sample was digested with endo H prior to electrophoresis and fluorography.

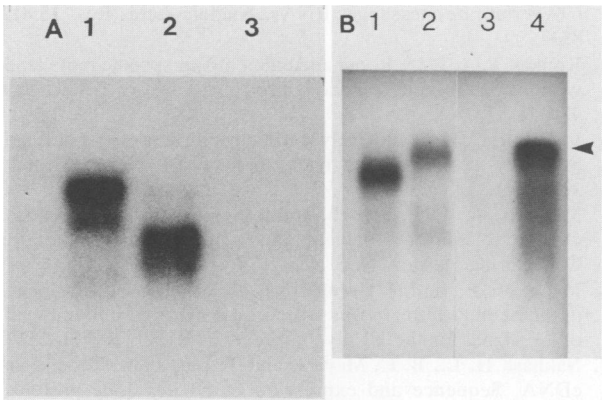

FIG. 4. (A) Analysis of E1-specific RNA from transfected COS cells. Total cellular RNA was isolated from transfected COS cells, separated on denaturing agarose gels, transferred to nitrocellulose, and hybridized to a ^{32}P-labeled 30-mer complementary to E1. Lanes: 1, pSVL-3'E2E1-transfected COS cell RNA; 2, pSVL-E1-transfected COS cell RNA; 3, pSVL-transfected COS cell RNA. (B) S1 nuclease digestion of RNA from transfected COS cells. Total cellular RNA was isolated from transfected COS cells, hybridized to ^{32}P-labeled antisense RNA from p3'E2E1, and digested with S1 nuclease; samples were then denatured, separated on a denaturing agarose gel, transferred to Hybond N (Amersham), and autoradiographed. Lanes: 1, RNA from pSVL-E1-transfected cells; 2, RNA from pSVL-3'E2E1-transfected cells; 3, RNA from vector-transfected cells; 4, antisense p3'E2E1 RNA (ca. 1,800 nts). The arrowhead indicates full-length 3'E2E1 RNA in lane 4.

pSVL-3'E2E1-specific RNAs were present in comparable amounts in the transfected cells (Fig. 4A). In addition to the major bands, there appeared to be minor amounts of other hybridizing RNA species (Fig. 4A, lanes 1 and 2). S1 nuclease protection analysis was performed to determine the sizes of the protected transcripts in transfected cells. Full-length transcripts were indeed synthesized in pSVL-3'E2E1-transfected cells (Fig. 4B, lane 2), and the protected RNA species in pSVL-E1-containing cells was only slightly smaller (Fig. 4B, lane 1). These results suggest that the undetectable level of pSVL-E1 protein in transfected COS cells is not due to decreased mRNA synthesis or aberrant RNA splicing.

DISCUSSION

The mechanism of processing of RV structural proteins E1 and E2 is largely unknown, despite the recently published sequence data (3, 26). Oker-Blom (18) has shown that the gene order on the 24S RNA is NH_2-C-E2-E1-COOH and, using synchronized translation experiments, that E2 is synthesized before E1. The question of whether E1 has its own independently functioning signal sequence has been addressed by expressing two different E1 constructs, one of which contains the putative signal, in cell-free translation systems and in COS cells. Plasmid p3'E2E1 contains the E1-coding region preceded by a highly hydrophobic stretch of amino acids with all the characteristics of a classical signal peptide, as described by von Heijne (27). These include a hydrophobic core, alanine at position −3, glycine at position −1, and proline at position −4 relative to the cleavage site (25, 27). In construct pE1, on the contrary, the putative signal is lacking, and translation starts at a methionine immediately adjacent to the E1 gene.

RNA from p3'E2E1 was translated in the presence of canine microsomes and produced a luminally sequestered polypeptide slightly larger than that produced by RNA translated in the absence of membranes. Detergent-labile protection from added trypsin was also afforded to this protein by the microsomes. The membrane topology of E1 predicted from these results implies that the 27 hydrophobic residues in the carboxyl end of E1 function as the transmembrane anchor (3, 26). The 13 residues exposed to the cytoplasm would therefore be the only part of E1 vulnerable to exogenously added protease, resulting in no significant decrease in the size of the protein. In contrast, the E1 protein lacking the signal peptide was not translocated into microsomes and subsequently was completely sensitive to trypsin.

The expression of pSVL-E1 and pSVL-3'E2E1 in COS cells further supports data for an independently functioning signal peptide for E1. The E1-specific protein synthesized in COS cells with pSVL-3'E2E1 is translocated into the ER and glycosylated. The endo H sensitivity of this protein indicates that the glycan moieties are of the high-mannose type (11). In RV virions, the E1 glycoproteins are of the endo H-resistant type (data not shown), indicating that E1 from pSVL-3'E2E1-transfected COS cells contains different oligosaccharide structures. Immunoprecipitates from COS cells transfected with pSVL-E1 do not contain detectable levels of E1, despite the presence of large amounts of intact E1-specific RNA present. This fact, however, does not rule out the possibility that mRNA from this plasmid is not translated as efficiently as that from pSVL-3'E2E1 in COS cells. However, inefficient translation of pSVL-E1 mRNA is unlikely in light of the fact that the ribosome-binding site 5'GCCTATGGCATGG3' in this construct is in fairly good agreement with the consensus sequence GCCGCC$^{A}_{G}$C-CATGG for eucaryotic ribosome-binding sites (12). Note that at the two most important positions, −3 and +4 (relative to ATG), homology is observed in the pSVL-E1 insert. Also, SP6 transcripts from pE1 efficiently direct translation in a cell-free system, thus proving that the ribosome-binding site common to pE1 and pSVL-E1 is functional. Clarke et al. (D. M. Clarke, T. W. Loo, H. McDonald, and S. Gillam, Gene, in press) have shown that unglycosylated E1 expressed in vivo is still antigenic, ruling out the possibility that the pSVL-E1 protein is not immunoprecipitated due to a lack of N-linked glycans.

Newly synthesized membrane proteins that fail to undergo translocation have been shown to be extremely unstable in the cytoplasm (7, 9, 23). The most likely reason for not detecting E1 protein in COS cells transfected with pSVL-E1 is rapid proteolysis of the untranslocated protein.

The strategy for the expression of RV structural proteins is similar to that in alphaviruses, such as Semliki Forest virus and Sindbis virus (6, 22), although RV does not contain a 6,000-dalton peptide between E2 and E1. The signal peptide of Semliki Forest virus E1 is contained in the C-terminal region of the 6,000-dalton peptide and is able to initiate translocation of E1 in the absence of E2, both in vitro and in vivo (16). Presumably, in RV E2-E1 processing, targeting of the E2-E1 precursor to the rough ER via the E2 signal sequence initiates the insertion of E2 into the ER lumen. The transfer of E2 is halted by the 19-residue transmembrane domain of E2, and translocation of E1 is initiated by the E1 signal peptide (Fig. 1B).

In summary, we have provided experimental evidence that a functional signal peptide for RV E1 glycoprotein exists within the C-terminal 69 amino acids of E2 and is likely the 20 residues preceding E1. This group of hydrophobic amino acids has been previously implicated as the signal peptide of E1 based on its conformity to features of well-known eu-

caryotic signal sequences (5, 17, 25, 26). We are currently engaged in the fine mapping of the E1 signal peptide by using mutants with deletions in the E2 carboxyl terminus.

ACKNOWLEDGMENTS

This work was supported jointly by grants from the Medical Research Council of Canada and the British Columbia Health Care Research Foundation. T.C.H. is a predoctoral trainee receiving support from the Medical Research Council of Canada in the biotechnology training program.

LITERATURE CITED

1. **Adams, G. A., and J. K. Rose.** 1985. Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport. Cell **41**:1007–1015.
2. **Blobel, G., and B. Dobberstein.** 1975. Transfer of proteins across membranes. 1. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J. Cell Biol. **67**:835–851.
3. **Clarke, D. M., T. W. Loo, I. Hui, P. Chong, and S. Gillam.** 1987. Nucleotide sequence and in vitro expression of rubella virus 24S subgenomic mRNA encoding the structural proteins E1, E2 and C. Nucleic Acids Res. **15**:3041–3057.
4. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
5. **Frey, T. K., I. D. Marr, M. L. Hemphill, and G. Dominguez.** 1986. Molecular cloning and sequencing of the region of the rubella virus genome coding for glycoprotein E1. Virology **154**: 228–232.
6. **Garoff, H., C. Kondor-Koch, and H. Riedel.** 1982. Structure and assembly of alphaviruses. Curr. Top. Microbiol. Immunol. **99**: 1–50.
7. **Gething, M. J., and J. Sambrook.** 1982. Construction of influenza haemagglutinin genes that code for intracellular and secreted forms of the protein. Nature (London) **300**:598–603.
8. **Hovi, T., and A. Vaheri.** 1970. Infectivity and some physiochemical characteristics of rubella virus ribonucleic acids. Virology **42**:1–8.
9. **Jabbar, M. A., and D. P. Nayak.** 1987. Signal processing, glycosylation, and secretion of mutant hemagglutinins of human influenza virus by *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7**: 1476–1485.
10. **Kalkkinen, N., C. Oker-Blom, and R. F. Pettersson.** 1984. Three genes code for rubella virus structural proteins E1, E2a, E2b, and C. J. Gen. Virol. **65**:1549–1557.
11. **Kornfeld, R., and S. Kornfeld.** 1985. Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. **54**:631–664.
12. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15**:8125–8132.
13. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**: 4753–4757.
14. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Melancon, P., and H. Garoff.** 1986. Reinitiation of translocation in the Semliki Forest virus structural polyprotein: identification of the signal for the E1 glycoprotein. EMBO J. **5**:1551–1560.
17. **Nakhasi, H. L., B. C. Meyers, and T. Liu.** 1986. Rubella virus cDNA. Sequence and expression of E1 envelope protein. J. Biol. Chem. **261**:16616–16621.
18. **Oker-Blom, C.** 1984. The gene order for rubella virus structural proteins is NH_2-C-E2-E1-COOH. J. Virol. **51**:354–358.
19. **Oker-Blom, C., N. Kalkkinen, L. Kääriäinen, and R. F. Pettersson.** 1983. Rubella virus contains one capsid protein and three envelope glycoproteins, E1, and E2a, and E2b. J. Virol. **46**:964–973.
20. **Oker-Blom, C., I. Ulmanen, L. Kääriäinen, and R. F. Pettersson.** 1984. Rubella virus 40S genome RNA specifies a 24S subgenomic mRNA that codes for a precursor to structural proteins. J. Virol. **49**:403–408.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Schlesinger, S., and M. J. Schlesinger.** 1986. The Togaviridae and Flaviviradae. Plenum Publishing Corp., New York.
23. **Sekikiwa, K., and C. J. Lai.** 1983. Defects in functional expression of an influenza virus hemagglutinin lacking the signal peptide sequences. Proc. Natl. Acad. Sci. USA **80**:3563–3567.
24. **Singer, S. J., P. A. Maher, and M. P. Yaffe.** 1987. On the transfer of integral proteins into membranes. Proc. Natl. Acad. Sci. USA **84**:1960–1964.
25. **Strauss, A. S., I. Boime, and G. Kreil.** 1986. Protein compartmentalization, p. 15–70. Springer-Verlag, New York.
26. **Vidgren, G., K. Takkinen, N. Kalkkinen, L. Kaarianen, and R. F. Pettersson.** 1987. Nucleotide sequence of the genes coding for the membrane glycoproteins E1 and E2 of rubella virus. J. Gen. Virol. **68**:2347–2357.
27. **von Heijne, G.** 1984. How signal sequences maintain cleavage specificity. J. Mol. Biol. **173**:243–251.
28. **Walter, P., and G. Blobel.** 1983. Preparation of microsomal membranes for cotranslational protein translocation. Methods Enzymol. **96**:84–93.