VIROLOGY 246, 74–82 (1998)
ARTICLE NO. VY989179

Proteolytic Processing of Rubella Virus Nonstructural Proteins

Jiansheng Yao,* Decheng Yang,* Pele Chong,† Dorothy Hwang,* Yuying Liang,* and Shirley Gillam*[1]

*Department of Pathology and Laboratory Medicine, Faculty of Medicine, University of British Columbia, 950 West 28th Avenue, Vancouver, British Columbia, Canada, V5Z 4H4; and †Connaught Research Institute, Willowdale, Ontario, Canada, M2R 3T4

Received November 21, 1997; returned to author for revision December 31, 1997; accepted April 7, 1998

The genomic RNA of rubella virus contains two long open reading frames (ORF), a 5′-proximal ORF that codes for the nonstructural proteins and a 3′-proximal ORF that encodes the structural proteins. The cDNA encoding the nonstructural protein ORF of the wild-type M33 strain of rubella virus has been obtained and sequenced. Comparison between the nonstructural proteins of the M33 and Therien strains of rubella virus revealed a 98% homology in nucleotide sequence and 98.1% in deduced amino acid sequence. To examine the processing of rubella virus nonstructural protein, the complete nonstructural protein ORF was expressed in BHK cells using a pSFV expression vector. Three nonstructural protein products (p200, p150, and p90) with molecular weights of 200, 150, and 90 kDa were identified using antisera raised against synthetic peptides corresponding to regions of the nonstructural proteins. p200 is the polyprotein precursor, while p150 and p90 are the cleavage products. Site-directed mutagenesis of the Cys-1151 residue (one of the catalytic dyad residues of the viral protease) and of the Gly-1300 residue (the viral protease cleavage site) abrogated protease activity and p200 precursor cleavage, respectively. Coexpression of mutant constructs in BHK cells indicated that rubella virus protease can function both in cis and in trans.

Key Words: Rubella virus; nonstructural proteins; polyprotein processing; expression.

INTRODUCTION

Rubella virus (RV) is an enveloped, positive-strand RNA virus in the family Togaviridae (Francki *et al.,* 1991). It is the sole member of the genus *Rubivirus* and bears striking similarities to the alphaviruses, such as Semliki Forest virus (SFV) and Sindbis virus (SV), in terms of genome organization and structural protein expression strategy (Strauss and Strauss, 1994). However, unlike alphaviruses, RV infection in cultured cells is characterized by a relatively long latency period, slow replication, and limited cytopathology and does not shut off host cell macromolecular synthesis (Hemphill *et al.,* 1988). Problems of slow replication and low virus yield have hampered progress regarding the molecular details of RV replication.

In the RV virion the icosahedral nucleocapsid is composed of a single copy of the 40S genomic RNA and multiple copies of the capsid protein (C, 33 kDa). This is surrounded by a lipid bilayer into which two viral glycoproteins, E1 (57 kDa) and E2 (42–47 kDa), are embedded (Oker-Blom *et al.,* 1983). In addition to the 40S genomic RNA, RV-infected cells contain a 24S subgenomic RNA that is derived from the 3′ end of the 40S genomic RNA (Oker-Blom *et al.,* 1984). Upon infection, the genomic RNA serves as both the mRNA for synthesis of the viral nonstructural proteins and the template for synthesis of complementary minus-strand RNA. Newly synthesized minus-strand RNA in turn functions as the template for both synthesis of plus-strand genomic RNAs and transcription of subgenomic RNAs, which are translated to produce the three viral structural proteins, C, E2, and E1 (Oker-Blom *et al.,* 1984).

Although the translation and processing of the structural proteins of RV have been studied in detail, little is known about the translation and processing of the nonstructural proteins (ns proteins). Bowden and Westaway (1984) demonstrated that there were four prominent species of RV-specific ns proteins (200, 150, 87, and 75 kDa); the 200-kDa species is the polyprotein precursor and the 150-kDa protein was derived from the 200-kDa precursor. Recently, when a cDNA containing the coding region of the ns proteins of RV was introduced into a vaccinia expression vector, three RV-specific proteins with molecular sizes of 200, 150, and 90 kDa were observed (Marr *et al.,* 1994).

Comparative amino acid sequence analyses of proteins involved in the replication of positive-strand RNA viruses, together with functional analyses of a few of these proteins, have allowed identification of amino acid sequence motifs for RNA replicases, RNA helicases, proteases, and RNA methyltransferases. The sequence motifs identified by these studies have allowed localization of sequences in RV ns polyprotein that are responsible for each of these functions. The p150 contains the amino-terminal sequences of the polyprotein and the

[1] To whom reprint requests should be addressed. Fax: +1 (604) 875-2496.

0042-6822/98 $25.00

proposed methyltransferase and protease sequences at its amino and carboxy terminus, respectively (Marr *et al.,* 1994). The p90 contains the predicted helicase and RNA polymerase sequences at its amino- and carboxy-terminal regions, respectively (Dominguez *et al.,* 1990). Experimental evidence for the protease (Chen *et al.,* 1996) and helicase (Gros and Wengler, 1996) activities of the proposed regions has been obtained. It has been demonstrated that RV p150 possesses a papain-like cysteine protease activity (Marr *et al.,* 1994), with the catalytic dyad residues of Cys-1151 and His-1272 and the protease cleavage site within Gly-1300–Gly-1301 (Chen *et al.,* 1996).

Understanding the mechanism of RNA virus replication is of prime significance because of its central role in the pathogenicity of a large group of viruses. Despite its importance, very little is known about the details of this replication process in higher eukaryotic cells. In order to study the replication mechanism of RV, we have cloned and expressed RV ns protein genes in BHK cells. We report here our studies on the cloning, sequencing, expression, and proteolytic cleavage of RV ns protein. A pSFV-1 eukaryotic expression system, based on the SFV replicon (Liljestrom and Garoff, 1991), was used in this study. Identification of RV ns proteins was carried out by using rabbit antisera raised against six synthetic peptides corresponding to different regions of RV ns proteins. To investigate the ability of RV protease to function in cis or in trans, several mutant constructs were generated. Coexpression studies indicate that RV protease can function both in cis and in trans.

RESULTS

Sequence analysis and comparison between M33 and Therien strains

Sequence analysis of the RV ns protein genes (M33 strain) revealed a long open reading frame (ORF) of 6345 nucleotides with an AUG initiation codon beginning at nucleotide 41 and terminating at nucleotide 6385 with an ochre (UAA) stop codon which was followed by two further in-frame ochre stop codons at nucleotides 6410 and 6425, respectively (EMBL Accession No. X72393RVM33NP). This ORF differs from that reported for the Therien strain, in which the 5′-proximal ORF is 6615 nucleotides in length and terminates at position 6656 with an opal codon (UGA) which is followed 12 nucleotides downstream by a second inframe opal codon (Dominguez *et al.,* 1990). This ORF overlaps with the 3′-proximal ORF beginning at nucleotide 6507 (Dominguez *et al.,* 1990). This difference was due to a sequencing error in the Therien strain (omission of two nucleotides (CG) at position 6292). With these two additional nucleotides, the 5′-proximal ORF would use different termination codons than originally deduced in the Therien strain (Dominguez *et al.,* 1990). The corrected termination codons would be identical to those of the M33 strain reported here. This sequencing error in the Therien strain has been corrected (Pugachev *et al.,* 1997).

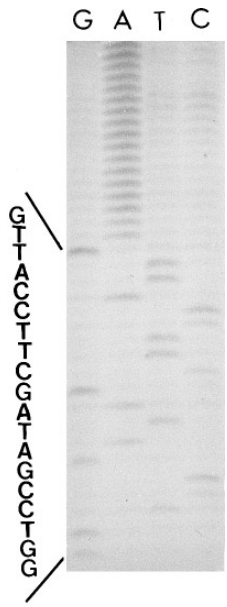

FIG. 1. The 5′-terminal sequence of RV genomic RNA. RV virion RNA was used as template in the first-strand cDNA synthesis. The 3′-terminally poly(A)-tailed cDNA was used as a template in PCR amplification as described under Materials and Methods. The DNA sequence downstream from the A tract is shown at the right of the gel.

Comparison of the ns protein ORF between M33 and Therien strains revealed a 98% homology in nucleotide sequence and 98.1% in deduced amino acid sequence. We have also determined the 5′ end of the genomic RNA of M33 strain by poly(A) tailing of cDNA followed by PCR amplification. The sequence obtained from the amplified product is shown in Fig. 1. The sequence GTTAC, downstream of poly(A), is complementary to CAATG of the 5′ terminus of the cDNA of the genomic RNA. Therefore, the RV M33 strain also has the same 5′-terminal sequence as that reported in the Therien strain (Wang *et al.,* 1994).

Expression of RV ns proteins in transfected BHK cells

The pSFV-1 eukaryotic expression vector (Liljestrom and Garoff, 1991) is a novel DNA expression system based on the SFV replicon, which has a broad host range and a high level of expression efficiency. To study the processing of the RV ns protein, a cDNA containing the RV ns protein ORF was inserted into the *Sma*I site of the pSFV-1 vector. Recombinant plasmid (pSFV/RVnsp) (Fig. 2) linearized at the *Spe*I site was used for the synthesis of RNA transcripts using phage SP6 RNA polymerase. BHK cells were transfected with synthesized RNA transcripts by electroporation. After 24 h posttransfection, BHK cells were labeled with [^{35}S]methionine/cysteine and ^{35}S-labeled cellular lysates were immunoprecipitated with antisera raised against synthetic peptides corresponding to the amino acid sequence of the RV ns

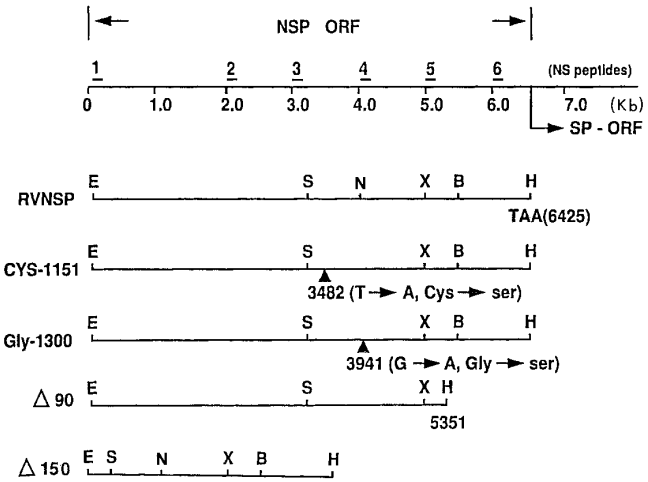

FIG. 2. Schematic representation of pSFV/RVnsp mutants and regions of the NSP ORF corresponding to the synthetic peptides. The RV sequences within the NSP ORF are shown with a number line in kilobases. Numbers below line in the diagram represent nucleotides in RV genome. The mutation resulting in a Cys to Ser or a Gly to Ser conversion in the pSFV/RVnsp construct is indicated (▲). The deletion mutants Δ90 and Δ150 were constructed as described under Materials and Methods. The regions corresponding to the synthetic peptides (NS1, NS2, NS3, NS4, NS5, and NS6) are indicated above the diagram representing the RV genome. S, *Sph*I; X, *Xba*I; B, *Bgl*II; N, *Nco*I; H, *Hin*dIII; and E, *Eco*RI.

protein (Fig. 2). Three ORF-specific proteins, p200, p150, and p90 (molecular masses of 200, 150, and 90 kDa, respectively), were identified using peptide antisera (NS1, NS2, NS4, NS5, and NS6) (Fig. 3A). In BHK cells transfected with the pSFV-1 vector, no specific protein species with molecular sizes of 200, 150, and 90 kDa were immunoprecipitated by any of the antisera (Fig. 3A, lanes 1, 3, 5, 7, 9, and 11). The failure of peptide antiserum NS3 to immunoprecipitate any specific ns protein products may be due to the low affinity of this antiserum. Aside from this failure the recognition of p200 by all of the antisera indicates that p200 is the polyprotein precursor, and p150 and p90 are the cleavage products of p200. Since p150 was immunoprecipitated by antisera NS1, NS2, and NS4, whereas p90 was recognized by antisera NS5 and NS6, the order of the cleavage products is NH_2-150-90-COOH. This is consistent with the finding of Forng and Frey (1995). As peptide antisera NS1 and NS5 were found to give higher titers than the other antisera, they were used for the studies reported below.

To further examine the processing of RV ns protein, the transfected BHK cells were pulse-labeled for 40 min and chased for the indicated time periods prior to lysis. After a 40 min of pulse labeling, p150 (Fig. 3B, lane 3) and p90 (Fig. 3B, lane 4) were detected (Fig. 5B, lanes 3 and 4), indicating that the cleavage of p200 is quite efficient. Chased with excess unlabeled methionine and cysteine, the intensities of bands corresponding to p150 and p90 increased, whereas that for p200 decreased after a 120-min chase (Fig. 3B, lanes 5 to 10). The other protein species with molecular weights of 170, 130, and 110 kDa (indicated as dots) could be the degraded products of ns proteins or nonspecific host proteins coprecipitated with the ns protein.

trans processing of ns protein

According to the prediction of Gorbalenya *et al.* (1991), RV protease belongs to the family of Main or M-protease that exhibits trans activity and mediates viral RNA replication and expression. However, it has been shown that when part of the RV ns protein ORF containing the protease domain and its cleavage site (nucleotides 2515 to 4560) was expressed in BHK cells by using a Sindbis virus vector (Bredenbeek and Rice, 1992), no trans cleavage activity was detected in co-expression studies (Chen *et al.,* 1996). To determine whether the lack of trans cleavage activity is due to the deletion at either the 5′ or the 3′ end of p200, we have

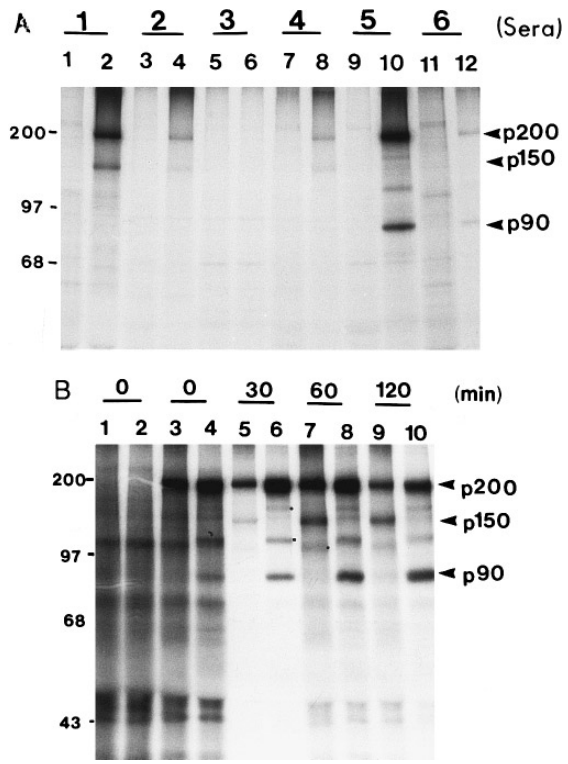

FIG. 3. Expression of RVnsp ORF in pSFV/RVnsp-transfected BHK cells. (A) BHK cells were transfected with a pSFV-1 vector (lanes 1, 3, 5, 7, 9, and 11) or pSFV/RVnsp (lanes 2, 4, 6, 8, 10, and 12). Transfected cells were labeled with [^{35}S]methionine/cysteine for 60 min and cellular lysates were immunoprecipitated with peptide antisera (NS1 to NS6). Positions of nsp ORF-specific products (200, 150, and 90 kDa) are marked on the right. Molecular weight standards (in kDa) are indicated on the left. (B) BHK cells were transfected with a pSFV vector (lanes 1 and 2) or with pSFV/RVnsp (lanes 3 to 10), pulsed labeled with [^{35}S]methionine/cysteine for 40 min, and chased for the indicated time periods. Cellular lysates were immunoprecipitated with antisera NS1 (lanes 1, 3, 5, 7, and 9) or NS5 (lanes 2, 4, 6, 8, and 10). RV-specific ns proteins are indicated on the right. Molecular weight standards (in kDa) are indicated on the left.

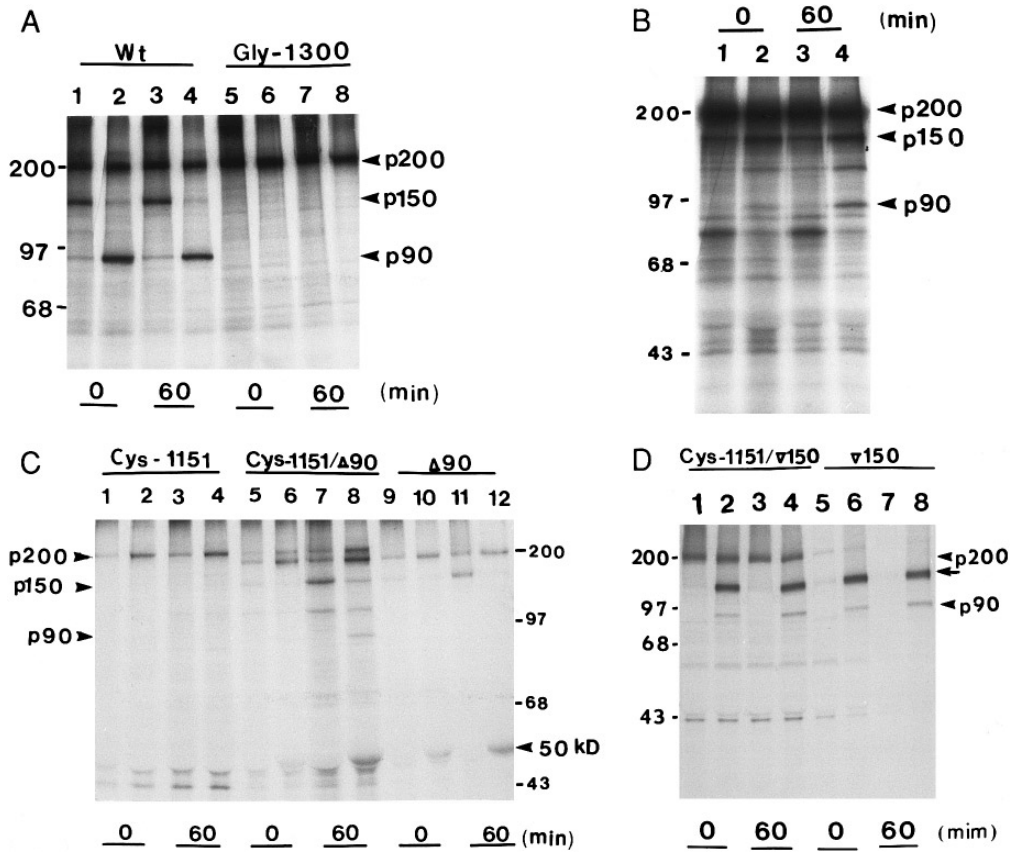

FIG. 4. Analysis of trans activity of RV ns proteins. (A) BHK cells were transfected with pSFV/RVnsp (wt) or with mutant construct pSFV/Gly-1300. Transfected BHK cells were labeled for 40 min and chased for 60 min. Cellular lysates were immunoprecipitated with antisera NS1 (lanes 1, 3, 5, and 7) or NS5 (lanes 2, 4, 6, and 8). RV-specific ns proteins are indicated on the right. (B) BHK cells were cotransfected with pSFV/Cys-1151 and pSFV/Gly-1300 constructs. Transfected BHK cells were labeled for 40 min and chased for 60 min. Cellular lysates were immunoprecipitated with antisera NS1 (lanes 1 and 3) or NS5 (lanes 2 and 4). RV-specific ns proteins are indicated on the right. Molecular weight standards (in kDa) are indicated on the left. (C) BHK cells were transfected with individual mutant constructs (pSFV/Cys-1151 or pSFV/Δ90) or both constructs. Transfected cells were pulse labeled for 40 min and chased for 60 min. Cellular lysates were immunoprecipitated with antisera NS1 (lanes 1, 3, 5, 7, 9, and 11) or NS5 (lanes 2, 4, 6, 8, 10, and 12). RV-specific ns proteins are indicated on the left. Molecular weight standards (in kDa) are indicated on the right. (D) BHK cells were transfected with individual mutant constructs (pSFV/Cys-1151 or pSFV/Δ150) or both constructs. Transfected cells were pulse labeled for 40 min and chased for 60 min. Cellular lysates were immunoprecipitated with antisera NS1 (lanes 1, 3, 5, and 7) or NS5 (lanes 2, 4, 6, and 8). RV-specific ns proteins are indicated on the right. The polyprotein of Δ150 mutant (130 kDa) is indicated by an arrow. Molecular weight standards (in kDa) are indicated on the left.

constructed four mutants in ns protein ORF. pSFV/Cys-1151, a protease-inactivated mutant carrying a cysteine to serine mutation at residue 1151; pSFV/Gly-1300, a cleavage site mutant carrying a glycine to serine mutation at residue 1300; pSFV/Δ90, a deletion mutant in which part of the coding region of p90 at the 3′ end of p200 was deleted; and pSFV/Δ150, a deletion mutant in which 918 amino acid residues (nucleotides 39 to 2794) at the 5′ end of p200 were deleted (Fig. 2).

As expected, changing glycine to serine at residue 1300 resulted in the complete abolition of cleavage; no cleavage product was detected after a 60-min chase (Fig. 4A, lanes 7 and 8). This cleavage-blocked mutant (pSFV/Gly-1300) was used as a source of protease in cotransfection experiments with protease-inactivated mutant (pSFV/Cys-1151) which provide p200 as a source of substrate. After a 40-min pulse labeling, p200 precursor and its cleavage products (p150 and p90) were observed (Fig. 4B, lanes 1 and 2), and the intensity of p150 and p90 bands increased after a 60-min chase (Fig. 4B, lanes 3 and 4). Since no cleavage products were detected when BHK cells were transfected alone with either pSFV/Cys-1151 (protease-inactivated mutant) (Fig. 4C, lanes 1 to 4) or pSFV/Gly-1300 (cleavage-blocked mutant) (Fig. 4A, lanes 5 to 8), the detection of cleavage products in the cotransfection experiments indicates that trans processing had occurred. To further confirm the trans activity of RV protease, we used a 3′-end deletion mutant (pSFV/Δ90) to serve as a source of protease in cotransfection experiments. After a 40-min pulse labeling, cleavage of this deletion precursor was observed (Fig. 4C, lanes 9 and 10). Production of the cleavage products (150 and 50 kDa) increased after a 60-min chase (Fig. 4C, Δ90, lanes 11 and 12). In the cotransfec-

tion experiments, after a 60-min chase, p90 was clearly visible (Fig. 4C, Cys-1151/Δ90, lane 8). These results clearly demonstrated that RV protease has trans activity, although the processing in trans was slower than that of the cis processing. To determine whether the 5′-end region upstream of the protease domain is necessary for the trans cleavage activity, we carried out the coexpression studies using a 5′-end deletion mutant (pSFV/Δ150) as a source of protease and protease-inactivated mutant (pSFV/Cys-1151) as a source of substrate. After a 40-min pulse labeling, cleavage of the 130-kDa deletion precursor (Fig. 4D, arrow) to p90 was observed (Fig. 4D, lane 6). Due to the lack of antiserum specific to the 40-kDa cleavage product, only p90 was immunoprecipitated by NS5 antiserum. The intensity of the p90 band increased after a 60-min chase (Fig. 4D, lane 8). However, in the cotransfection experiments, no trans cleavage activity was detected, as no p150 cleavage product was immunoprecipitated by NS1 antiserum (Fig. 4D, lanes 1 and 3). The p90 product in the coexpression experiments was the cleavage product from deletion precursor (130 kDa) (Fig. 4D, lanes 2 and 4). This finding indicates that the trans cleavage activity of RV protease may require the native p150.

Processing of RV ns protein in RV-infected Vero cells

Nakhasi *et al.* (1988) have shown that treatment of cells with actinomycin D either before infection or during the eclipse period of RV infection, but not during later periods, decreases the final yield of virus. These results indicate that during the viral infection, host cell DNA directs the synthesis of cellular factor(s) essential for viral replication. When the synthesis of this cellular factor(s) is terminated at an early stage of viral infection by actinomycin D, viral replication is impaired. To determine the processing of the RV ns protein at different times in the viral replication cycle and the association of host protein(s) with the replication complexes, RV-infected Vero cells were labeled for 3 h with [^{35}S]methionine/cysteine at 24 or 48 h postinfection (pi), followed by a 3-h chase with excess unlabeled methionine and cysteine. The monolayers were lysed with buffer containing 1% NP-40 (without SDS) and immunoprecipitation was performed as described by Wahlberg *et al.* (1989) using NP-40 as a detergent. At 24 h pi, p200 was the dominant ns protein species observed (Fig. 5A), while at 48 h pi, p90 was the major ns protein species (Fig. 5B). The intensity of the p90 band was much greater than that of the p150 band and the p200 band decreased substantially compared to that at 24 h pi (Fig. 5B). These results indicate that RV ns protein processing is relatively slow and it is possible that early in infection, the cleavage kinetics might differ from those in late infection and that the processing of p200 may be highly regulated during the infectious cycle.

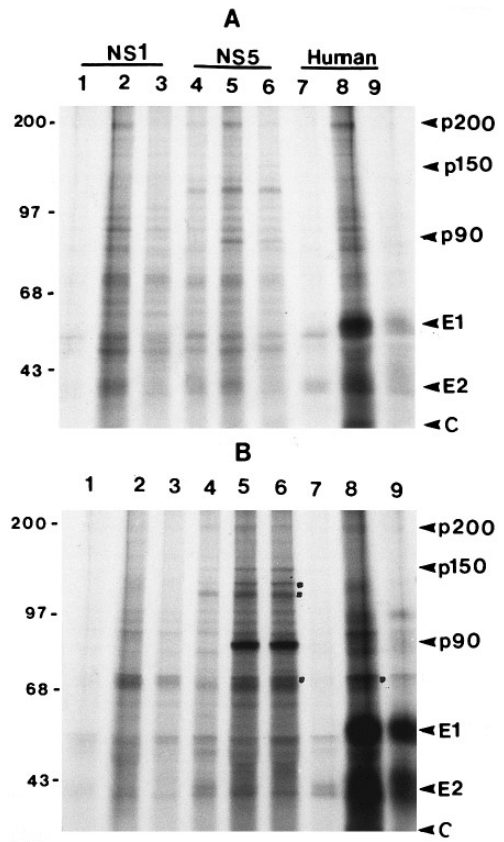

FIG. 5. Immunoprecipitation of RV ns proteins in infected cells. Monolayers of Vero cells mock-infected (lanes 1, 4 and 7) or RV-infected (lanes 2, 3, 5, 6, 8, and 9) were labeled for 3 h at 24 h pi (A) or 48 h pi (B) and chased for 3 h (lanes 3, 6, and 9). Monolayers were lysed with lysis buffer containing 1% NP-40, and cellular lysates were immunoprecipitated with antisera NS1 (lanes 1 to 3), NS5 (lanes 4 to 6), or human anti-RV serum (lanes 7 to 9). The position of RV-specific ns proteins and the molecular weight standards (in kDa) are indicated, respectively, on the right and the left.

In our earlier studies, we have observed that our antibodies coprecipitate other protein species besides p200, p150, and p90. In this study, in the absence of SDS, besides p200, p150, and p90, protein species with molecular sizes of 140, 130, and 70 kDa were co-immunoprecipitated by antiserum NS5 (Fig. 5B, lanes 5 and 6, dots). It is interesting that human anti-RV serum also co-immunoprecipitated p200, p90, and the 70-kDa protein, apart from the RV structural proteins (Fig. 5, human lane 8). It is likely that the 140- and 70-kDa protein species were host proteins coprecipitated by antisera. At present we do not know if these proteins are involved in RV replication.

The compound L-*trans*-epoxysuccinyl-L-leucylamido (4-guanidinobutane) (E-64) is a specific inhibitor of thiol proteases, including the papain family of cysteine proteases (Kleina and Grubman, 1992). To examine the effect of E-64 on the processing of p200, RV-infected Vero cells (48 h postinfection) were treated with E-64 or E-64C (an analog of E-64, *trans*-epoxysuccinyl-L-leucylamido-3-

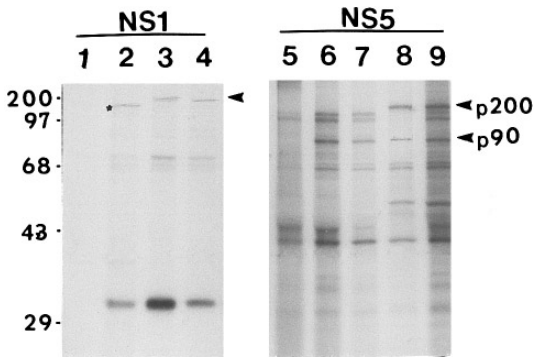

FIG. 6. Immunoprecipitation of cytoplasmic extracts from RV-infected Vero cells in the presence or the absence of E-64 and E-64C. Monolayers of Vero cells mock-infected (lanes 1 and 5) or RV-infected (lanes 2 to 4 and 6 to 9) were treated with E-64 (lanes 1, 3, 5, and 8) or E-64C (lanes 4 and 9) for 90 min at 48 h pi and labeled with [^{35}S]methionine/cysteine for 2.5 h. After a 2.5-h chase, monolayers were lysed, and cellular lysates were immunoprecipitated with antisera NS1 (lanes 1 to 4) or NS5 (lanes 5 to 9). The position of RV-specific ns proteins is indicated on the right. The molecular weight standards (in kDa) are indicated on the left. E-64 and E-64C were dissolved in dimethyl sulfoxide. The concentration of dimethyl sulfoxide in the incubation was 3% (except in lane 6, in which no dimethyl sulfoxide was added).

methylbutaine) at a concentration of 0.3 mg/ml for 90 min, and labelled with [^{35}S]methionine/cysteine for 2.5 h. After a 2.5-h chase with excess unlabeled methionine/cysteine, the monolayers were lysed and cellular lysates immunoprecipitated with antisera. (E-64 and E-64C were present in the medium during the pulse–chase period). In control infected cells, p200 was not detected (Fig. 6, lanes 2, 6, and 7). While in treated cells, p200 was still present after the chase period (Fig. 6, lanes 3, 4, 8, and 9), and processing of p200 was only partially inhibited, as p90 cleavage product was detected in the treated cells (Fig. 6, lanes 8 and 9). It appears that RV assembly was not affected by the treatment (data not shown). We could not use a higher concentration of E-64 or E-64C in our experiments (the reported range is between 0.1 and 2.0 mg/ml) (Kleina and Grubman, 1992), due to the limited solubility of E-64 and E-64C in aqueous solution and the inhibition of dimethyl sulfoxide on RV replication (data not shown). In spite of this limitation, our results provide additional evidence that RV protease belongs to the papain family.

DISCUSSION

We have cloned and sequenced the complete ns protein genes of the RV M33 strain. Sequence analysis of the coding region of ns protein genes revealed a long ORF with an AUG initiation codon at nucleotide 41 and three in-frame ochre stop codons at nucleotides 6386, 6410, and 6425, respectively. Comparison between the ns proteins of the M33 and Therien strains of RV showed a 98% homology in nucleotide sequence and 98.1% in deduced amino acid sequence. Three ns protein ORF-specific products with molecular weights of 200, 150, and 90 kDa were identified using antisera raised against synthetic peptides corresponding to regions of ns proteins of the RV M33 strain (Fig. 3). Our results are consistent with the reported work of Forng and Frey (1995).

Two classes of viral papain-like protease have been assigned on the basis of their ability to function in cis or trans (Gorbalenya *et al.,* 1991). Main protease cleaves the viral polyprotein both in trans and in cis at multiple sites within polyprotein and contains a new conserved domain (the x domain, a novel conserved domain of unknown function), whereas leader protease exhibits cis activity and cleaves at single site (Gorbalenya *et al.,* 1991). By comparative sequence analysis, Gorbalenya *et al.,* (1991) proposed that RV protease is a Main protease and bears pronounced similarity to cellular proteases. Our studies demonstrated that RV protease is a Main protease that can function both in cis and in trans (Fig. 4), consistent with the prediction of Gorbalenya *et al.* (1991). The failure to detect trans cleavage by the RV protease when it was expressed as a truncated form (from residues 2515 to 4560) (Chen *et al.,* 1996) could be due to conformational changes induced by the deletion or the requirement of intact p150 for the expression of trans cleavage activity.

In general, most ns proteins of viruses arise from the processing of a completed precursor. The protein species responsible for the trans cleavage is not known. We presume that p150 once formed can also act in trans as a protease, since the 3′-end mutant polyprotein (pSFV/Δ90) containing part of the p90 was found to be able to act in trans (Fig. 4C), and this trans cleavage activity was abolished when the 5′-end region including the x domain (residues 1 to 918) was deleted (Fig. 4D). It is tempting to speculate that the x domain might be involved in the regulation of polyprotein processing. Further studies of the capabilities of proteins individually expressed to function as protease and of RV mutants that are unable to cleave the polyprotein will be instructive.

MATERIALS AND METHODS

Cells and viruses

BHK-21 and Vero cells were grown in Eagle's minimum essential medium (MEM) supplemented with 10% fetal bovine serum (FBS), penicillin (100 units/ml), and streptomycin (50 μg/ml). RV (M33 strain) was propagated in Vero cells and isolated from the culture supernatant as described previously (Clarke *et al.,* 1987).

Cloning and sequencing of RV ns protein genes

Intracellular total RNA from RV-infected Vero cells was isolated at 72 h postinfection (m.o.i. = 0.1). The isolated cellular RNAs were used in first-strand cDNA synthesis with random deoxyhexamer as primers (Rice *et al.,* 1985).

The cDNA fragments were amplified using the polymerase chain reaction (PCR) (Erlich, 1989) with *Taq* DNA polymerase and synthetic oligonucleotides designed according to the sequence of RV Therien strain (Dominguez *et al.,* 1990). Nine PCR fragments isolated by agarose gel electrophoresis were treated with T4 DNA polymerase and inserted into the *Sma*I site of the M13mp19 vector. The orientations of the inserts in positive clones were identified by complementary tests using M13 supernatant phages (Messing, 1983).

For sequencing, single-stranded DNA in both orientations was prepared for each clone. Universal M13 primer and synthetic oligonucleotides corresponding to appropriate regions of ns protein genes were used in sequencing. In some cases, sequencing reactions were carried out at 75°C using *Taq* DNA polymerase to avoid problems caused by the secondary structure of the template. Sequencing was performed on an ABI automated DNA sequencer (Applied Biosystems Inc). Both the dye-primer and dye-terminator kits were used in the sequencing reaction. For some regions of DNAs, manual sequencing (Sanger *et al.,* 1977) was used to confirm ambiguities in the sequence. Sequence analysis was performed using DNA Strider program (Marck, 1988). The complete nucleotide sequence for nonstructural protein genes of RV (M33 strain) is deposited in Gene Bank under Accession No. X72393RVM33NP.

Determination of the 5′-terminal sequence of virion RNA

RV particles were isolated from culture supernatants (50 ml) of RV-infected Vero cells by the conventional polyethylene glycol precipitation (Bowden and Westaway, 1984). Pelleted RV particles were resuspended in 0.5 ml TNE buffer (10 mM Tris–HCl, pH 7.5; 150 mM NaCl; 1 mM EDTA), and virion RNA was extracted with Trizol reagent (Gibco/BRL). First-strand cDNA synthesis was carried out at 42°C for 1 h in a 20-μl reaction mixture containing virion RNA, 0.5 μg random hexamer (Promega), 200 units of superscript RT RNase H^- reverse transcriptase, and 40 units of RNasin (Promega) in a buffer provided by the manufacturer (Gibco/BRL). The synthesized cDNA was purified using a Qiaquick PCR purification kit (Qiagen Inc) and then 3′-terminally tailed with dATP by using terminal deoxynucleotidyltransferase (Boehringer Mannheim Biochemicals). The poly(A)-tailed cDNA was extracted with phenol/chloroform, precipitated with ammonium acetate, and used as a template for the following amplification. A cDNA fragment of about 80 nucleotides containing a poly(A) tail was amplified in a 50-μl PCR reaction mixture containing an upstream oligonucleotide with a poly(T) tract (5′-GAATTCAAGT$_{17}$-3′), a downstream oligonucleotide (5′-CCTCATCTAGGAGTTTC-3′), complementary to the RV genome at nucleotides 46 to 61), and 2 units of Vent (exo^-) DNA polymerase (New England Biolabs) in a buffer provided by the manufacturer. The amplification consisted of one cycle of 1 min at 95°C; 2 cycles of 30 s at 50°C, 1 min at 70°C; and 33 cycles of 20 s at 95°C, 20 s at 50°C, and 40 s at 70°C. The amplified fragments were purified with a Qiaquick PCR purification kit and cloned into the *Sma*I site of the pSU18 vector (Martinez *et al.,* 1988). Three clones were sequenced using the downstream oligonucleotide that was used in the PCR reaction as a primer in dideoxy sequencing (Sanger *et al.,* 1977). The complementary DNA sequence obtained is shown in Fig. 1.

Peptide synthesis and rabbit immunization

Six synthetic peptides within ns protein (NS1, residues 1–36; NS2, residues 614–648; NS3, residues 1003–1033; NS4, residues 1272–1307; NS5, residues 1598–1637; and NS6, residues 1900–1933) (Fig. 2) were synthesized in an ABI 430A automated peptide synthesizer using solid-phase methods (Merrifield, 1969). Synthesized peptides were cleaved from the resin by treatment with hydrogen fluoride and purified by reverse-phase high-pressure liquid chromatography using a Vydac C4 column. The purity of all peptide preparations exceeded 95%. For each peptide, amino acid analysis was performed on a Waters Pico-tag system and found to be in good agreement with the theoretical composition.

To prepare peptide-specific antisera, New Zealand white rabbits (two rabbits for each peptide) were immunized intramuscularly with each peptide (500 μg per rabbit) emulsified in Freund's complete adjuvant. On days 21, 52, 82, 112, and 142, each rabbit received booster injections with the same amount of respective immunogen in incomplete Freund's adjuvant. Rabbit sera were collected 2 weeks after the final booster injection.

RNA transcription and transfection

The cDNA encoding the full-length RV ns protein-ORF (Fig. 2) was inserted into the *Sma*I site of the pSFV-1 vector (BRL). This construct was named pSFV/RVnsp. The plasmid pSFV/RVnsp was linearized at the *Spe*I site, and the linearized DNA was used as the template for *in vitro* transcription by standard protocols (Promega Corp) using each nucleoside triphosphate and mG(5)pp(5)G at final concentrations of 1 mM each. BHK-21 cells (0.5 ml, 10^7 cells/ml) were transfected with transcribed RNA by electroporation (Liljestrom and Garoff, 1991). The electroporated cells were suspended in 10 ml MEM containing 10% FBS and seeded in a 6-well culture plate.

Metabolic labeling

Transfected BHK cells. BHK cells at 24 h posttransfection were washed once with D-MEM lacking methionine and cysteine. Then 5 ml of D-MEM containing 3% dialyzed FBS but lacking methionine and cysteine was added. After 30 min incubation at 37°C, ^{35}S-protein la-

beling mix (1175 Ci/mmol, NEN) was added to the medium at a concentration of 100 μCi/ml and the incubation was continued for 40 min. In pulse–chase experiments, at the end of labeling period, the medium containing the radiolabel was removed, and D-MEM containing the normal concentrations of methionine and cysteine was added. Following the completion of the labeling and chase period, the monolayers were washed three times with PBS and scraped into 1 ml PBS. The cells were spun down and resuspended in TNE buffer (10 mM Tris–HCl, pH 7.5; 150 mM NaCl; 1 mM EDTA; and 10 mM phenylmethylsulfonyl fluoride) containing 1% SDS.

RV-infected Vero cells. Subconfluent monolayers of Vero cells grown in 60-mm petri dishes were mock-infected or infected with RV at a m.o.i. of 0.5 PFU/cell. At 24 or 48 h postinfection, monolayers were washed with PBS twice and covered with D-MEM lacking methionine and cysteine. Following incubation for 45 min at 37°C, ^{35}S-protein labeling mix (1175 Ci/mmol, NEN) was added to the medium at the concentration of 100 μCi/ml and then further incubated for an additional 3 h. For pulse–chase experiments, the protocol was the same as described above in transfected BHK cells.

Immunoprecipitation

Rabbit anti-peptide serum was preincubated with Protein A–Sepharose (Pharmacia) for 2 h in binding buffer (100 mM Tris–HCl (pH 7.4), 400 mM NaCl) with constant mixing at room temperature. The serum-coated beads were washed twice with binding buffer and once in lysis buffer (25 mM Tris–HCl (pH 7.4), 100 mM NaCl, 1 mM EDTA, 1% Nonidet P-40). ^{35}S-Labeled cell lysate was mixed with the coated beads overnight at 4°C in lysis buffer. Beads were washed once with lysis buffer; twice with wash buffer (25 mM triethanolamine; 172 mM NaCl; 1% deoxycholate; 0.1% SDS; 1 mM EDTA; pH 7.4); once with 10 mM Tris–HCl (pH 7.4); and once with water. Immunocomplexes were dissociated from the Protein A–Sepharose by boiling in 100 mM sodium citrate (pH 5.5), 0.15% SDS for 5 min, vortexing, and pelleting the beads by centrifugation.

To determine the coprecipitation of host proteins with RV ns proteins, immunoprecipitation of cellular lysates from RV-infected Vero cells was performed as described by Wahlberg *et al.,* (1989). Briefly, at the end of the pulse or chase period, monolayers were washed twice with PBS and lysed with buffer containing 50 mM Tris–HCl (pH 7.4), 150 mM NaCl, 2 mM EDTA, 1% Nonidet P-40 (NP-40), and 74 μM antipain dihydrochloride (Boehringer-Mannheim). Peptide antisera-coated Protein A–Sepharose was washed with lysis buffer and incubated with ^{35}S-labeled cellular lysates overnight at 4°C. Beads were washed twice with washing buffer containing 10 mM Tris–HCl (pH 7.4), 150 mM NaCl, 2 mM EDTA, and 0.2% NP-40, and once with 10 mM Tris–HCl (pH 7.4). Immunocomplexes were eluted as described above.

Construction of mutants

Two mutant plasmids (pSFV/Cys-1151 and pSFV/Gly-1300) were constructed by using oligonucleotide-directed mutagenesis on a single-stranded uracil-containing template (Kunkel *et al.,* 1987). The fragment (*Sph*I/*Xba*I) from pSFV/RVnsp (Fig. 2) was isolated and subcloned into the phage vector M13mp18 which had been digested with *Sph*I and *Xba*I restriction enzymes. Mutagenic oligonucleotides were 5′-GACCCAAACACCAGCTGGCTCCGCGCC-3′ (nucleotides 3494 to 34508) and 5′CTGTCTCGGGGCAGCGGCACTTGTGCC-3′ (nucleotides 3929 to 3958), respectively, for Cys-1151 to serine and Gly-1300 to serine mutations (mutated bases are underlined). After verification of the introduced mutation by DNA sequencing, the *Sph*I/*Xba*I fragment containing the introduced mutation was recloned into the pSFV/RVnsp plasmid (minus the original *Sph*I/*Xba*I fragment). pSFV/Cys-1151 carried the cysteine to serine mutation at residue 1151, and pSFV/Gly-1300 had the glycine residue at 1300 changed to serine (Fig. 2).

The deletion mutant (pSFV/Δ90) was constructed by deletion of the *Bgl*II/*Hin*dIII fragment (nucleotides 5352 to 6494) from pSFV/RVnsp (Fig. 2). The stop codon cassette in the pSFV-1 vector was used in this construct. To create a deletion mutant (pSFV/Δ150, deletion of nucleotides 39 to 2794), the *Nco*I fragment (nucleotides 39 to 4020) was first removed from the pSFV/RVnsp (Fig. 2) and then ligated with the *Nhe*I/*Nco*I fragment (nucleotides 2800 to 4020) containing a *Nco*I site at the 5′ end of the *Nhe*I site. For production of the *Nhe*I/*Nco*I fragment, the upstream primer was 5′-ATTCCCATGGTCGCGCTAGCGCC-3′ (*Nco*I and *Nhe*I sites are underlined) and the downstream primer was 5′-GTAGGTGGCGGCGTTCTTGAT-3′ (nucleotides 4217 to 4237). The amplified PCR product was digested with *Nco*I prior to ligation.

ACKNOWLEDGMENTS

This work was supported by a grant from the Medical Research Council of Canada. S.G. is an investigator of the British Columbia Children's Hospital Foundation.

REFERENCES

Bredenbeek, P. J., and Rice, C. M. (1992). Animal RNA virus expression systems. *Semin. Virol.* 3, 297–310.

Bowden, D. S., and Westaway, E. G. (1984). Rubella virus structural and non-structural proteins. *J. Gen. Virol.* 65, 933–943.

Chen, J. P., Strauss, J. H., Strauss, E. G., and Frey, T. K. (1996). Characterization of the rubella virus nonstructural protease domain and its cleavage site. *J. Virol.* 70, 4707–4713.

Clarke, D. M., Loo, T. W., Hui, I., Chong, P., and Gillam, S. (1987). Nucleotide sequence and *in vivo* expression of rubella virus 24S subgenomic mRNA encoding the structural proteins E1, E2 and C. *Nucleic Acids Res.* 15, 3041–3057.

Dominguez, G., Wang, C. Y., and Frey, T. K. (1990). Sequence of the

genome RNA of rubella virus: Evidence for genetic rearrangement during Togavirus evolution. *Virology* 177, 225–238.

Erlich, H. A. (Ed.). (1989). "PCR: Principles and Applications for DNA Amplification." Stockton Press, New York.

Forng, R. Y., and Frey, T. K. (1995). Identification of the rubella virus nonstructural proteins. *Virology* 206, 843–853.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L., and Brown, F. (Eds.). (1991). "Classification and Nomenclature of Viruses." Fifth report of the International Committee on Taxonomy of Viruses. Archives of Virology, Suppl. 2, Springer–Verlag, Vienna.

Gorbalenya, A. E., Koonin, E. V., and Lai, M. M-C (1991). Putative papain-related thiol proteases of positive-strand RNA viruses: Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, α-, and coronaviruses. *FEBS Lett.* 288, 201–205.

Gros, C., and Wengler, G. (1996). Identification of an RNA-stimulated NTPase in the predicted helicase sequence of the rubella virus nonstructural polyprotein. *Virology* 217, 367–372.

Hemphill, M. L., Forng, R. Y., Abernathy, E. S., and Frey, T. K. (1988). Time-course of virus-specific macromolecular synthesis in rubella virus infected Vero cells. *Virology* 162, 65–75.

Kleina, L. G., and Grubman, M. J. (1992). Antiviral effects of a thiol protease inhibitor on foot-and-mouth disease virus. *J. Virol.* 66, 7168–7175.

Kunkel, T., Roberts, J. D., and Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection methods. *Methods Enzymol.* 154, 367–387.

Liljestrom, P., and Garoff, H. (1991). A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Bio/Technology* 9, 1356–1361.

Marck, C. (1988). DNA Strider: A "C" program for the fast analysis of DNA and protein sequence on the Apple Macintosh family computers. *Nucleic Acids Res.* 16, 1829–1836.

Marr, L. D., Wang, C. Y., and Frey, T. K. (1994). Expression of the rubella virus nonstructural protein ORF and demonstration of proteolytic processing. *Virology* 198, 586–592.

Martinez, E., Bartolome, B., and de la Cruz, F. (1988). pACYC184-derived cloning vectors containing the multiple cloning site and lac Zα reporter gene of pUC8/9 and pUC18/19 plasmids. *Gene* 68, 159–162.

Merrifield, R. B. (1969). Solid phase peptide synthesis. *Adv. Enzymol.* 32, 221–296.

Messing, J. (1983). New M13 vector for cloning. *Methods Enzymology* 101, 20–78.

Nakhasi, H. L., Zheng, D., Hewlett, I. K., and Liu, T. Y. (1988). Rubella virus replication: Effect of interferons and actinomycin D. *Virus Res.* 10, 1–15.

Oker-Blom, C., Kalkkinen, N., Kääriäinen, L., and Pettersson, R. F. (1983). Rubella virus contains one capsid protein and three envelope glycoproteins, E1, E2a, and E2b. *J. Virol.* 46, 964–973.

Oker-Blom, C., Ulmanen, I., Kääriäinen, L., and Pettersson, R. F. (1984). Rubella virus 40S genomic RNA specifies a 24S subgenomic mRNA that codes for a precursor to structural proteins. *J. Virol.* 49, 403–408.

Pugachev, K. V., Abernathy, E. S., and Frey, T. K. (1997). Genomic sequence of the RA 27/3 vaccine strain of rubella virus. *Arch. Virol.* 142, 1165–1180.

Rice, C. M., Lenches, E. M., Eddy, S. R., Shin, S. J., Sheets, R. L., and Strauss, J. H. (1985). Nucleotide sequence of yellow fever virus: Implification for flavirirus gene expression and evolution. *Science* 229, 726–733.

Sanger, F., Nicklem, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitor. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.

Strauss, J. H., and Strauss, E. G. (1994). The alphaviruses: Gene expression, replication, evolution. *Microbiol. Rev.* 58, 491–562.

Wahlberg, J. W., Boere, W. A. M., and Garoff, H. (1989). The heterodimeric association between the membrane proteins of Semliki Forest virus changes its sensitivity to low pH during virus maturation. *J. Virol.* 63, 4991–4997.

Wang, C. Y., Dominguez, G., and Frey, T. (1994). Construction of rubella virus genomic-length cDNA clones and synthesis of infectious RNA transcripts. *J. Virol.* 68, 3550–3557.