# Complementation of a Deletion in the Rubella Virus P150 Nonstructural Protein by the Viral Capsid Protein

Wen-Pin Tzeng and Teryl K. Frey\*

Department of Biology, Georgia State University, Atlanta, Georgia 30302-4010

Received 20 March 2003/Accepted 9 June 2003

Rubella virus (RUB) replicons with an in-frame deletion of 507 nucleotides between two *Not*I sites in the P150 nonstructural protein ( $\Delta Not$ I) do not replicate (as detected by expression of a reporter gene encoded by the replicon) but can be amplified by wild-type helper virus (Tzeng et al., Virology 289:63-73, 2001). Surprisingly, virus with  $\Delta Not$ I was viable, and it was hypothesized that this was due to complementation of the *Not*I deletion by one of the virion structural protein genes. Introduction of the capsid (C) protein gene into  $\Delta Not$ I-containing replicons as an in-frame fusion with a reporter gene or cotransfection with both  $\Delta Not$ I replicon, confirming the hypothesis that the C gene was the structural protein gene responsible for complementation and demonstrating that complementation. Mutations that prevented translation of the C protein while minimally disturbing the C gene sequence abrogated complementation, while synonymous codon mutations that changed the C gene sequence without affecting the amino acid sequence at the 5' end of the C gene had no effect on complementation, indicating that the C protein, not the C gene RNA, was the moiety responsible for complementation occurred at a basic step in the virus replication cycle, because  $\Delta Not$ I replicons failed to accumulate detectable virus-specific RNA.

Rubella virus (RUB) is the sole member of the genus Rubivirus in the family Togaviridae (for a review, see reference 7). The RUB virion consists of a genomic, single-stranded RNA enclosed in a quasispherical capsid composed of multiple copies of the viral capsid protein, C, which is in turn surrounded by a lipid bilayer envelope in which are embedded two virus glycoproteins, E1 and E2. The RUB genome is 9,762 uncleotides (nt) in length, of positive-polarity, and contains two long open reading frames (ORFs). The 5'-proximal ORF, or nonstructural protein ORF (NS-ORF), is translated from the genome RNA into a 240-kDa precursor that is proteolytically cleaved at a single site by a virus-encoded protease into two products: an N-terminal product of 150 kDa (P150) and a C-terminal product of 90 kDa (P90). The 3'-proximal ORF, or structural protein ORF (SP-ORF), which is translated from a subgenomic (SG) RNA, encodes the virion proteins in the order 5'-C-E2-E1-3'; processing of these proteins is mediated by the cellular enzyme signal endopeptidase.

The RUB NS proteins function in viral RNA replication. From predictions based on computer alignment with sequences from other viruses, P150 contains (from N terminus to C terminus) a methyltransferase domain, a Y domain, a proline hinge domain, an X domain, and a protease domain that catalyzes the cleavage of the NS precursor; P90 contains helicase and RNA-dependent RNA polymerase (RDRP) domains (12). Of these, the activities of the protease and helicase domains have been confirmed experimentally (3, 8, 15, 16, 19, 20, 22, 34, 35). While it has been hypothesized that the X domain functions in *trans* cleavage mediated by the protease (15), putative functions for the Y and proline hinge domains have not been proposed.

Recently, we developed a RUB replicon (RUBrep) in which the SP-ORF was replaced with a reporter gene, such as that coding for chloramphenical acetyltransferase (CAT) or green fluorescent protein (GFP) (32). While the replicon RNA itself is replicated and the reporter gene is expressed only in the initially transfected cell, in the presence of wild-type helper virus, the replicon is amplified, packaged, and spread to other cells. In this context, the replicon resembles naturally occurring RUB defective-interfering RNAs that maintain the NS-ORF but contain large deletions within the SP-ORF. None of the RUBrep/GFP constructs with deletions in the NS-ORF was capable of replication (as detected by GFP expression) as expected. However, of these deletion mutants, only a construct with an in-frame deletion between two NotI sites at nt 1685 and 2192 of the genome (thus designated  $\Delta NotI$ ) could be amplified by wild-type helper virus. This deletion encompassed amino acid residues 551 to 720 of P150, a region between the Y and proline hinge domains. This finding indicated that the NS-ORF is required in cis for amplification by wild-type helper virus as well as for self-replication, probably due to a *cis* preference by the viral replicase proteins for the RNA template from which they were translated (1, 17, 32). We hypothesized that only the *Not*I region of the NS proteins, which has no known or suggested function based on computer alignments, could be complemented by wild-type helper virus, but then were surprised to find that when  $\Delta NotI$  was introduced into an infectious cDNA clone (Robo402), a viable virus was produced that replicated to within threefold of wild-type virus (32). This indicated that another viral gene could complement deletions of the NotI region. In this report, we show that the C protein can complement deletions of the NotI region of the P150 NS protein.

<sup>\*</sup> Corresponding author. Mailing address: Georgia State University, Biology Department, P.O. Box 4010, Atlanta, GA 30302-4010. Phone and Fax: (404) 651-3105. E-mail: tfrey@gsu.edu.

### MATERIALS AND METHODS

**Recombinant DNA methods.** Recombinant DNA manipulations were performed basically as described by Sambrook et al. (27) with minor modifications. *Escherichia coli* JM109 was used as the bacterial host. Restriction enzymes and T4 DNA ligase were obtained from New England BioLabs (Beverly, Mass.) or Roche Molecular Biochemicals (Indianapolis, Ind.) and used essentially as recommended by the manufacturers. Standard PCR mixtures contained 400 ng of each oligonucleotide primer, 20 ng of linearized plasmid template, 200  $\mu$ M each deoxynucleotide triphosphate, and 5 U of Ex *Taq* DNA polymerase (PanVera/ TaKaRa, Madison, Wis.) in 1× buffer provided by the manufacturer in a total volume of 50  $\mu$ l. The cycling protocol was 20 s at 98°C, 20 s at 50°C, and 1 to 3 min at 70°C for 35 cycles followed by one cycle of 10 min at 72°C.

Generation of constructs and site-directed mutagenesis. The infectious genomic cDNA clones Robo402, Robo402- $\Delta NotI$ , and Robo402/IRES (internal ribosome entry site construct), all of which use pBR322 as a backbone plasmid, were described previously (25, 26, 33). The SP6 RNA polymerase promoter-cDNA insert from each of these clones was transferred to pUC18 to create Robo502, Robo502- $\Delta NotI$ , and Robo502/IRES, respectively. An additional genomic cDNA construct, Robo503, was generated, which could be linearized prior to in vitro transcription with either *EcoRI* or *SpeI* (this was necessary for constructs, RUBrep/GFP and RUBrep/GFP- $\Delta NotI$ , were described previously (32, 33). For use in this study, the SP6 RNA polymerase promoter-cDNA insert in each of these constructs was transferred to pUC18.

To create the mutated constructs used in this study, both PCR amplification with mutagenic oligonucleotide primer or primers followed by restriction fragment replacement with the amplification product into the parent vector and a three-round asymmetric PCR-three-fragment ligation strategy (33) were employed. Constructs with fusions of the C or E2 genes with the GFP gene or C, E2, or E1 with 3'-terminal deletions fused with GFP were initiated with mutagenic oligonucleotides that placed an XbaI site following the 3' end of the gene or the deletion within the gene that allowed in-frame fusion with an XbaI site at the 5<sup>th</sup> end of the GFP cassette. Constructs that contained progressive deletions of the C gene from the 5' end were initiated with mutagenic oligonucleotides that placed an XbaI or SpeI site and an ATG initiation codon upstream from the nucleotides following the deletion. RUBrep constructs expressing a fusion protein between the complete C gene of Sindbis virus (SIN) or the SIN C gene with 3' deletions and GFP were initiated with an upstream oligonucleotide containing EcoRI and SpeI sites and the 5' 21 nt of the SIN C gene and downstream oligonucleotides that contained an XbaI site followed by 15 to 16 nt complementary to the sequence of the SIN C gene upstream from the deletion site, allowing for introduction into SpeI-XbaI-digested RUBrep/GFP; XhoI-linearized pTE5'2J (10) was used as the template. A series of mutations in the 5' nine codons of the C gene were initiated with mutagenic oligonucleotides containing the desired mutation or mutations with 8  $\sim$ 12 nt on either side of the mutation. Mutations created in RUBrep/GFP were transferred to RUBrep/CAT (33). To generate replicons containing the NotI deletion, the HindIII-BglII fragment from RUBrep/GFP-ΔNotI (containing the SP6 RNA polymerase promoter and the 5' end of the RUB genome through nt 5355, a region containing the NotI deletion) was used to replace the corresponding fragment in the RUBrep/GFP or RUBrep/CAT construct. Plasmids expressing cassettes containing various regions of the RUB NS-ORF and SP-ORF as well as the SIN SP-ORF under control of the human cytomegalovirus (CMV) immediate-early promoter were created with VR1012 vector (Vical, Inc., San Diego, Calif.). Each cassette was amplified by PCR with an upstream oligonucleotide that contained a restriction site appropriate for introduction into the multiple cloning site of VR1012, an ATG for initiation of translation if necessary, and the 5' 15 nt from the region to be expressed. The downstream oligonucleotide contained an appropriate restriction site, the complement of a termination codon if necessary, and nucleotides complementary to the 3' 15 nt from the region to be expressed.

In vitro transcription and transfection. Robo, RUBrep, and VR1012 constructs were purified on CsCl isopycnic density gradients prior to use. Robo502 and its derivatives were linearized with *Eco*RI, while Robo503 and its derivatives were linearized with *SpeI* prior to in vitro transcription, which was carried out as previously described (33). The transcription reaction mixtures were used directly for cell transfection without DNase treatment or phenol-chloroform extraction. Vero and BHK-21 cells were transfected with Lipofectamine 2000 as previously described (33). GFP expression was monitored by direct examination of the transfected monolayer with a Zeiss Axioplan microscope with epifluorescence capability. For CAT activity assay, Vero cells were transfected with transcripts from replicon-CAT- $\Delta NorI$  constructs. At 4 days posttransfection, cells were lysed and used for CAT activity assay basically as described by Seed and Sheen (28). Replicon-specific RNAs species synthesized in transfected cells were analyzed by Northern assay with a NorthernMax-Gly kit (Ambion, Houston, Tex.) as previously described with <sup>32</sup>P-labeled, nick-translated pGEM/GFP as a probe (32).

## RESULTS

C gene complements NotI deletions in RUB replicons in cis. We showed previously that transcripts from a replicon construct expressing the GFP gene, RUBrep/GFP, produced GFP following transfection of Vero cells, while transcripts from a construct with the NotI region of the NS-ORF deleted (RUBrep/GFP- $\Delta NotI$ ) did not (32). However, in the presence of coinfecting wild-type helper RUB or cotransfection with transcripts from the infectious cDNA clone, Robo402, the GFP signal was amplified for both constructs. Unexpectedly, transcripts from a Robo402 construct containing the NotI deletion, Robo402- $\Delta Not$ I, yielded viable virus that replicated to within threefold of Robo402 virus. The only difference between Robo402-ANotI and RUBrep/GFP-ANotI is the SP-ORF. Therefore, we hypothesized that a structural gene or protein could complement the NotI region of the NS-ORF. allowing both replication of Robo402-ANotI virus and amplification of RUBrep/GFP- $\Delta NotI$  in the presence of wild-type helper virus. To this end, a series of replicons were constructed that contained in-frame fusions of the individual structural proteins with GFP. As shown in Fig. 1, while RUBrep/GFP-ΔNotI did not express GFP in transfected cells, RUBrep/C-E2-E1(1-9)-GFP- $\Delta NotI$  (which contains the SP-ORF through the first nine amino acids of the E1 gene fused to the GFP gene) and RUBrep/C-E2-GFP-ΔNotI (which contains the C and E2 genes fused with GFP gene) produced GFP, indicating that the structural proteins can complement the NotI deletion in the NS-ORF. Further analysis showed that RUBrep/C-E2(1-7)-GFP- $\Delta Not$ I, which contains the C gene through the first seven amino acids of the E2 gene fused to the GFP gene, and RUBrep/C-GFP- $\Delta Not$ I, which contains the C gene fused with GFP gene, did express GFP in transfected cells, while RUBrep/E2-GFP- $\Delta Not$ I, which contains the E2 gene fused with GFP gene, did not. When introduced into the wild-type replicon without the NotI deletion, RUBrep, all of these constructs expressed GFP. Thus, the moiety responsible for complementation of the NotI deletion was the C gene.

Next, to examine whether the complete C gene or only part would suffice to complement the NotI deletion, a series of C-GFP fusions that contained progressive deletions of the C gene from either the 5' or 3' end, or both, were generated. As shown in Table 1, when present in RUBrep- $\Delta Not$ I, none of the 5' deletions, including one that only deleted 8 codons, expressed GFP, and only the smallest 3' deletion, which deleted 23 codons (i.e., the E2 signal sequence at the C terminus of C) expressed GFP. In RUBrep, all of these constructs expressed GFP. Thus, the majority of the C gene was required for complementation. We also made a series of constructs expressing a fusion protein between the C gene of the alphavirus SIN, or 3'-truncated fragments of the SIN C gene, and GFP in both RUBrep and RUBrep- $\Delta Not$ I. As shown in Table 1, all of these RUBrep constructs expressed GFP, while none of these RUBrep- $\Delta Not$ I constructs expressed GFP, and thus the SIN C protein cannot substitute for the RUB C protein in complementing the NotI deletion.

# Α

| <u>Cassette</u>      | <b>Replication</b>                     | <u>Comple</u>                                 | <u>mentation</u>                                     |
|----------------------|--|---|--|
|                      |  | <u>Cis</u>                                    | <u>Trans</u>   |
|                      | GFP cassette in<br>RUBrep<br>(GFP/RNA) | GFP cassette in<br>RUBrep-∆NotI<br>_(GFP/RNA) | CAT cassette in RUBreı<br>+RUBrep/GFP-∆NotI<br>(GFP) |
| GFP/CAT              | +/+                                    | - /-  | -  |
| C-E2-E1(1-9)-GFP/CAT | 8 +/+                                  | +/+   | +  |
| C-E2-GFP/CAT         | +/+                                    | +/+   | +  |
| C-E2(1-7)-GFP/CAT    | +/+                                    | +/+   | +  |
| C-GFP/CAT            | +/+                                    | +/+   | +  |
| E2-GFP/CAT           | +/+                                    | - /-  | -  |

SP-ORF C IEZE

📙 E2 signal peptide 🛛 🗏 E1 signal peptide



FIG. 1. Mapping the moiety responsible for complementation. In panel A and Table 1 are shown the results of experiments with cassettes expressed by RUBrep or RUBrep-\DeltaNotI. The cassettes consisted of either a reporter gene (GFP or CAT) or a portion of the RUB or SIN SP-ORF fused in frame with GFP or CAT. The GFP version of each cassette was used to assess replication of the wild-type replicon, RUBrep (replication column), and of the replicon with the NotI deletion, RUBrep-ANotI (cis complementation column), containing the cassette. Replication or complementation was detected by both GFP expression using fluorescence microscopy of transfected cells (GFP) and virus-specific RNA production using Northern blot analysis (B [RNA]). trans complementation was assessed by cotransfecting cells with transcripts of the RUBrep construct containing the CAT version of the cassette and RUBrep/GFP- $\Delta Not$ I transcripts; replication of RUBrep/GFP- $\Delta Not$ I was detected by GFP expression. Panel A summarizes experiments with fusion proteins containing portions of the RUB SP-ORF. A schematic diagram of the ORF with the coding sequences for the C, E2, and E1 proteins as well as the E2 and E1 signal sequences (which remain attached to C and E2, respectively, following processing) is shown at the top of the panel and the portion of the ORF contained in each cassette, along with the reporter gene, is shown under this diagram. The C-E2-E1(1-9)-GFP/CAT cassette contains the complete C and E2 genes and the first nine amino acids of the E1 gene fused to the reporter gene, and the C-E2(1-7)-GFP/CAT cassette contains the complete C gene and the first seven amino acids of the E2 gene fused to the reporter gene. For the Northern blot shown in panel B, Vero cells were transfected with transcripts from RUBrep/GFP- $\Delta NotI$  (lane 1), RUBrep- $\Delta Not$  constructs expressing a series of C-GFP- $\Delta Not$  fusions that contained progressive 3'-terminal deletions of the C gene (lanes 2 to 11), RUBrep/C-GFP-ΔNotI (lane 12), RUBrep/C-E2-GFP-ΔNotI (lane 13), RUBrep/E2-GFP-ΔNotI (lane 14), RUBrep/C-E2-E1(1-9)-GFP- $\Delta Not$  I (lane 15), or RUBrep/C-E2(1-7)-GFP- $\Delta Not$  I (lane 16). Four days posttransfection, total RNA was extracted and analyzed by gel electrophoresis, blotting, and probing with <sup>32</sup>P-labeled pGEM-GFP DNA. The positions of migration of the replicon genomic (G) RNAs and SG RNAs (which vary in size) are denoted.

| TABLE 1. | Experiments with | fusion proteins | containing | terminally      |
|----------|------------------|-----------------|------------|-----------------|
| trur     | cated fragments  | of the RUB and  | SIN C gen  | es <sup>a</sup> |

| Cassette         Replication<br>(GFP/RNA)         cis (GFP/<br>RNA)         transpace           RUB C         3' deletions         (GFI)           C(1-277)-GFP/CAT $+/+$ $+/+$ $+/+$ C(1-277)-GFP/CAT $+/+$ $-/ +$ C(1-199)-GFP/CAT $+/+$ $-/ +$ C(1-151)-GFP/CAT $+/+$ $-/ +$ C(1-118)-GFP/CAT $+/+$ $-/ +$ C(1-50)-GFP/CAT $+/+$ $-/ +$ C(1-51)-GFP/CAT $+/+$ $-/ +$ C(1-18)-GFP/CAT $+/+$ $-/ -$ C(1-8)-GFP/CAT $+/+$ $-/ -$ S' deletions         C(9-88)-GFP/CAT $+/+$ $-/-$ C(9-300)-GFP/CAT $+/+$ $-/ -$ <th></th> <th>Deulisstien</th> <th>Complement</th> <th>ntation</th> |                    | Deulisstien | Complement        | ntation        |
|---|--------------------|-------------|-------------------|----------------|
| RUB C         3' deletions $C(1-277)$ -GFP/CAT       +/+ $++$ +/+ $C(1-199)$ -GFP/CAT       +/+ $++$ -/- $++$ -/- $C(1-151)$ -GFP/CAT       +/+ $++$ -/- $++$ -/- $C(1-18)$ -GFP/CAT       +/+ $++$ -/- $++$ -/- $C(1-58)$ -GFP/CAT       +/+ $-+$ -/- $C(1-50)$ -GFP/CAT       +/+ $-+$ -/- $C(1-50)$ -GFP/CAT       +/+ $$ -/- $C(1-44)$ -GFP/CAT       +/+ $$ -/- $C(1-18)$ -GFP/CAT       +/+ $$ -/- $C(1-8)$ -GFP/CAT       +/+ $$ -/- $$ -/- $$ -/- $$ -/- $$ -/- $$ -/- $$ -/- $$ -/- $$ -/-         <   | Cassette           | (GFP/RNA)   | cis (GFP/<br>RNA) | trans<br>(GFP) |
| 3' deletions<br>C(1-277)-GFP/CAT +/+ +/+ +<br>C(1-199)-GFP/CAT +/+ -/- +<br>C(1-151)-GFP/CAT +/+ -/- +<br>C(1-118)-GFP/CAT +/+ -/- +<br>C(1-88)-GFP/CAT +/+ -/- +<br>C(1-58)-GFP/CAT +/+ -/- +<br>C(1-50)-GFP/CAT +/+ -/- +<br>C(1-31)-GFP/CAT +/+ -/- +<br>C(1-31)-GFP/CAT +/+ -/- +<br>C(1-8)-GFP/CAT +/+ -/<br>C(1-8)-GFP/CAT +/+ -/<br>C(1-8)-GFP/CAT +/+ -/<br>C(1-8)-GFP/CAT +/+ -/   | RUB C              |             |                   |                |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | 3' deletions       |             |                   |                |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-277)-GFP/CAT   | +/+         | +/+               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-199)-GFP/CAT   | +/+         | -/-               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-151)-GFP/CAT   | +/+         | -/-               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-118)-GFP/CAT   | +/+         | -/-               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-88)-GFP/CAT    | +/+         | -/-               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-58)-GFP/CAT    | +/+         | -/-               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-50)-GFP/CAT    | +/+         | -/-               | +              |
| $\begin{array}{cccccccccccccccccccccccccccccccccccc$  | C(1-44)-GFP/CAT    | +/+         | -/-               | +              |
| C(1-18)-GFP/CAT +/+ -/<br>C(1-8)-GFP/CAT +/+ -/<br>5' deletions<br>C(9-88)-GFP/CAT +/+ -/<br>C(9-300)-GFP/CAT +/+ -/  | C(1-31)-GFP/CAT    | +/+         | -/-               | +              |
| C(1-8)-GFP/CAT +/+ -/<br>5' deletions<br>C(9-88)-GFP/CAT +/+ -/<br>C(9-300)-GFP/CAT +/+ -/  | C(1-18)-GFP/CAT    | +/+         | -/-               | -              |
| 5' deletions<br>C(9-88)-GFP/CAT +/+ -/<br>C(9-300)-GFP/CAT +/+ -/   | C(1-8)-GFP/CAT     | +/+         | -/-               | -              |
| C(9-88)-GFP/CAT +/+ -/<br>C(9-300)-GFP/CAT +/+ -/   | 5' deletions       |             |                   |                |
| C(9-300)-GFP/CAT +/+ -/   | C(9-88)-GFP/CAT    | +/+         | -/-               | _              |
|   | C(9-300)-GFP/CAT   | +/+         | _/_               | _              |
| C(18-88)-GFP/CAT +/+ -/   | C(18-88)-GFP/CAT   | +/+         | _/_               | _              |
| C(18-300)-GFP/CAT +/+ -/  | C(18-300)-GFP/CAT  | +/+         | _/_               | _              |
| C(19-300)-GFP/CAT +/+ -/  | C(19-300)-GFP/CAT  | +/+         | _/_               | _              |
| C(58-300)-GFP/CAT +/+ -/  | C(58-300)-GFP/CAT  | +/+         | _/_               | _              |
| C(88-300)-GFP/CAT +/+ -/  | C(88-300)-GFP/CAT  | +/+         | _/_               | _              |
| C(118-300)-GFP/CAT +/+ -/   | C(118-300)-GFP/CAT | +/+         | _/_               | _              |
| C(151-300)-GFP/CAT +/+ -/   | C(151-300)-GFP/CAT | +/+         | _/_               | _              |
| C(199-300)-GFP/CAT +/+ -/   | C(199-300)-GFP/CAT | +/+         | -/-               | _              |
| C(277-300)-GFP/CAT +/+ -/   | C(277-300)-GFP/CAT | +/+         | -/-               | _              |
|   |                    |             |                   |                |
| SIN C   | SIN C              |             |                   |                |
| C(1-264)-GFP/CAT +/+ -/   | C(1-264)-GFP/CAT   | +/+         | -/-               | -              |
| C(1-220)-GFP/CAT +/+ -/   | C(1-220)-GFP/CAT   | +/+         | -/-               | -              |
| C(1-160)-GFP/CAT +/+ -/   | C(1-160)-GFP/CAT   | +/+         | -/-               | -              |
| C(1-100)-GFP/CAT +/+ -/   | C(1-100)-GFP/CAT   | +/+         | -/-               | -              |
| C(1-40)-GFP/CAT +/+ -/  | C(1-40)-GFP/CAT    | +/+         | -/-               | _              |

<sup>*a*</sup> The designation for each construct gives the amino acid residues retained in the fragment expressed by that construct: the complete RUB C gene is C(1-300), while the complete SIN C gene is C(1-264). The replication results represent the GFP cassette in RUBrep, and the complementation results represent the GFP cassette in RUBrep- $\Delta Not$  (*cis*) and the CAT cassette in RUBrep + RUBrep/GFP- $\Delta Not$  (*trans*).

To test expression with a different reporter gene, the CAT gene was employed. As shown in Table 2, similar to the GFP constructs, RUBrep/CAT- $\Delta Not$ I did not express CAT; however, RUBrep/C-CAT- $\Delta Not$ I did. In contrast to the fusions of progressive 3' deletions of C to GFP, CAT fusions containing the N-terminal 88 codons of the C gene expressed CAT. This finding was likely related to the different virus production efficiencies following transfection of the versions of the infectious clone used to generate the CAT and GFP constructs. RUBrep/GFP and RUBrep/CAT were built by using Robo502 and Robo503, respectively. Robo503 has a SpeI 3' linearization in place of the EcoRI linearization site in Robo502; this substitution was necessary because the CAT gene contains an internal EcoRI site. Unexpectedly, while the transfection efficiencies of the Robo502 and Robo503 transcripts were similar, virus production was more rapid in Robo503-transfected cells than in Robo502-transfected cells such that by 4 days posttransfection, the difference in virus produced was 100-fold greater (data not shown). In the replicon context, this would

| TABLE 2 | 2. | GFP | and | CAT | activity | in | Vero | cells |
|---------|----|-----|-----|-----|----------|----|------|-------|
|         |    |     |     |     |          |    |      |       |

| Townstation                    | Activity         |                        |  |  |
|--------------------------------|------------------|------------------------|--|--|
| Iransiection                   | GFP <sup>a</sup> | CAT (cpm) <sup>b</sup> |  |  |
| Mock                           | _                | 374                    |  |  |
| RUBrep/CAT- $\Delta NotI$      | _                | 364                    |  |  |
| $C(1-8)$ -CAT- $\Delta NotI$   | _                | 350                    |  |  |
| $C(1-18)$ -CAT- $\Delta NotI$  | _                | 338                    |  |  |
| $C(1-31)$ -CAT- $\Delta NotI$  | —                | 328                    |  |  |
| $C(1-44)$ -CAT- $\Delta NotI$  | —                | 319                    |  |  |
| $C(1-50)$ -CAT- $\Delta NotI$  | —                | 352                    |  |  |
| $C(1-58)$ -CAT- $\Delta NotI$  | —                | 285                    |  |  |
| $C(1-88)$ -CAT- $\Delta NotI$  | —                | 16,185                 |  |  |
| $C(1-151)$ -CAT- $\Delta NotI$ | —                | 15,243                 |  |  |
| $C(1-199)$ -CAT- $\Delta NotI$ | —                | 287                    |  |  |
| $C(1-277)$ -CAT- $\Delta NotI$ | +                | 4,689                  |  |  |
| C-CAT- $\Delta NotI$           | +                | 2,324                  |  |  |
| C-E2-CAT- $\Delta Not$ I       | +                | 2,451                  |  |  |

<sup>*a*</sup> Vero cells were transfected with transcripts from GFP-Δ*Not*I. At 4 days posttransfection, GFP expression was monitored by direct examination of the transfected monolayer with a Zeiss Axioplan microscope with epifluorescence capability.

<sup>b</sup> Vero cells were transfected with transcripts from replicon-CAT-Δ*Not*I. At 4 days posttransfection, cells were lysed and assayed for CAT activity. CAT activity is given as the amount of [<sup>3</sup>H]acetyl coenzyme A partitioning into the aqueous phase following a reaction in the presence of chloramphenicol (32). The activity given is the average from three independent experiments. The low activity expressed by C(1-199)-CAT-ΔNotI is likely due to formation of aggregates by this fusion protein.

lead to greater production of the C protein-reporter gene fusion proteins. When the RUBrep/C-CAT- $\Delta Not$ I deletion series was tested in BHK cells, similar results were obtained as in Vero cells: namely that RUBrep/CAT- $\Delta Not$ I or RUBrep/C-CAT- $\Delta Not$ I fusions containing the 5' 58 codons did not express CAT, while RUBrep/C-CAT fusions containing 88 or more codons of the C gene expressed GFP (data not shown). Thus, the complementation phenomenon was not cell specific.

Thus far in this study, replication of RUBrep- $\Delta Not$ I constructs was assayed by reporter gene expression. To determine if complementation by the C gene was actually operating on SG RNA synthesis and/or translation of the reporter gene from the SG RNA, viral RNA synthesis in cells transfected with transcripts from the series of RUBrep/GFP- $\Delta NotI$  constructs was assayed by Northern blotting of intracellular RNA extracted from transfected cells. As shown in Fig. 1B, in cells transfected with constructs that did not express GFP (i.e., RUBrep/GFP-ΔNotI, RUBrep/E2-GFP-ΔNotI, and RUBrep/ C-GFP- $\Delta NotI$  containing less than the 277 codons of the C gene), no viral RNA was detected, while in cells transfected with constructs that expressed GFP, RUBrep/C(1-277)-GFP- $\Delta NotI$ , RUBrep/C-GFP- $\Delta NotI$ , RUBrep/C-E2(1-9)-GFP-ΔNotI, RUBrep/C-E2-GFP-ΔNotI, and RUBrep/C-E2-E1(1-7)-GFP- $\Delta Not$ I, both genomic and SG RNAs were readily detectable. RUBrep containing all of these constructs synthesized both genomic and SG RNAs (data not shown). This result revealed that the defect in  $\Delta NotI$  mutants and complementation by the C gene was at a basic point in the virus replication cycle, prior to accumulation of detectable virus RNA, and was not at the level of SG RNA synthesis and/or reporter gene translation. In subsequent experiments, complementation was assayed both by reporter gene expression and synthesis of viral RNA.

The C gene complements the NotI deletions in trans. Since amplification of replicons by wild-type helper virus occurs in trans, we next conducted experiments to determine if the C gene expressed in trans could complement RUBrep/GFP- $\Delta Not I$ . These experiments were performed by cotransfecting cells with RUBrep/GFP- $\Delta Not$ I transcripts and transcripts from RUBrep/CAT constructs; complementation was detected by expression of GFP from RUBrep/GFP-ΔNotI. While RUBrep/ CAT was unable to complement RUBrep/GFP- $\Delta NotI$ , demonstrating that the NS proteins cannot complement the NotI deletion, RUBrep/C-CAT fusion constructs containing minimally the 5' 31 codons as well as RUBrep/C-E2-CAT exhibited complementation (Fig. 1 and Table 1). This result was consistent with the ability of these C-CAT constructs to complement in cis (Table 1), with the exception that only 31 codons of the C gene was necessary for trans complementation, while 88 codons was necessary for cis complementation. None of the C-CAT fusion constructs that contained 5' deletions of the C gene, including the shortest, which deleted only eight codons, was able to complement RUBrep/GFP- $\Delta NotI$ . Additionally, none of the RUBrep/SIN-C-CAT constructs complemented in trans (Table 1).

As an independent system with which to analyze trans complementation, cells were cotransfected with RUBrep/GFP- $\Delta NotI$ , and plasmid vectors expressing regions of the RUB SP-ORF under control of the CMV immediate-early promoter; as described above, complementation was detected by expression of GFP by RUBrep/GFP- $\Delta NotI$ . CMV vectors expressing the complete SP-ORF, the C gene, or minimally the 5' 88 codons of the C gene exhibited complementation, while vectors expressing E2 and E1, the 5' 58 codons or less of the C gene, or the SIN SP-ORF did not complement. Using the CMV vector that expressed the shortest 5' region of the C gene able to complement, CMV RUB-C(1-88), a series of progressive 5' deletions of the C gene were made; however, none of these CMV constructs, including the shortest 5' deletion of eight codons, was able to complement RUBrep/GFP- $\Delta NotI$ . These results were consistent with trans complementation using RUBrep/CAT constructs, with the exception that only the 5' 31 codons of the C gene was necessary for complementation by RUBrep/CAT, while the 5' 58 codons were necessary for complementation by the CMV vector. The CMV vector was also used to analyze whether the NS-ORF region of the RUB could complement the NotI deletion. As shown in Table 3, CMV vectors expressing the entire RUB NS-ORF, the P150 gene, the P90 gene, or the NotI region were used to cotransfect cells with RUBrep/GFP- $\Delta Not$ I transcript; no GFP expression was observed.

Determination of whether the C protein or the C gene RNA is the complementing factor. Since the defect in  $\Delta Not$ I replicons is early in the replication cycle, before the accumulation of detectable virus RNA, we initially hypothesized that the C gene RNA was the moiety responsible for complementation. The C gene RNA could be provided by either the genomic or SG RNA, and therefore complementation by a RUB construct in which the "junction" untranslated region (UTR) between the NS- and SP-ORFs was replaced with the encephalomyocarditis virus IRES (Robo502/IRES) was tested; Robo502/ IRES does not synthesize SG RNA (26). When the *Not*I deletion was made in Robo502/IRES, the resulting construct,

TABLE 3. Detection of *trans* complementation with a plasmid-CMV vector<sup>a</sup>

|                     | 01                    |
|---------------------|-----------------------|
| CMV construct       | trans Complementation |
| Structural proteins |                       |
| CMV-SP-ORF          | +                     |
| CMV-SPE2E1          | —                     |
| CMV-C               | +                     |
| 3' C deletions      |                       |
| CMV-C(1-277)        | +                     |
| CMV-C(1-199)        | +                     |
| CMV-C(1-151)        | +                     |
| CMV-C(1-118)        | +                     |
| CMV-C(1-88)         | +                     |
| CMV-C(1-58)         | —                     |
| CMV-C(1-50)         | —                     |
| 5' C deletions      |                       |
| CMV-C(9-88)         | —                     |
| CMV-C(18-88)        | —                     |
| CMV-C(31-88)        | —                     |
| CMV-C(44-88)        | —                     |
| NS proteins         |                       |
| CMV-NSP             | —                     |
| CMV-p150            | —                     |
| CMV-p90             | —                     |
| CMV-NotI-NotI       | —                     |
| SIN SP-ORF          |                       |
| CMV-SIN-SP-ORF      | –                     |
|                     |                       |

<sup>*a*</sup> A series of plasmid constructs were generated with which coding regions of the RUB or SIN genome were expressed under control of the human CMV immediate-early promoter. In addition to the complete RUB and SIN SP-ORFs, the RUB E2-E1 genes (SPE2-E1; with the signal peptide [SP] of E2 included to ensure correct expression), and deleted fragments of the RUB C gene (retained amino acid residues given in the construct designation), constructs expressing the complete RUB NS-ORF (P200), P150 gene, P90 gene, and *NotI-NotI* region of the P150 gene were also generated. Note that these constructs were not reporter gene fusions. Vero cells were cotransfected with individual plasmid-CMV constructs and RUBrep/GFP- $\Delta NotI$  transcripts, and replication of RUBrep/GFP- $\Delta NotI$  was detected by GFP expression.

Robo502/IRES- $\Delta Not$ I, yielded viable virus, indicating that the SG RNA was not necessary for complementation (Fig. 2). Robo502/IRES- $\Delta Not$ I virus replicated about threefold less well than did Robo502/IRES virus, the same as the differential between Robo502 and Robo502- $\Delta Not$ I virus. When cells were cotransfected with RUBrep/GFP- $\Delta Not$ I transcripts and transcripts from either Robo502/IRES or Robo502/IRES- $\Delta Not$ I, complementation occurred, indicating that SG RNA was not necessary for complementation in *trans*. In addition to the lack of an SG RNA, the IRES constructs also lack the 5' UTR of the SG RNA, and thus this RNA sequence is not required for complementation.

We next focused on the 5' nine codons of the C gene that are necessary for complementation by using the C-GFP construct that contains the complete C gene. As shown in Fig. 3, the sequence at the 5' end of the C gene contains two in-frame AUGs (designated AUG1 and AUG2) separated by seven codons. A series of mutations of these sequences, shown in Fig. 3, were made in the C-GFP fusion protein cassette, in which the complete C gene is fused with GFP, and the mutated cassettes were introduced into (i) RUBrep/GFP to ascertain replication of the wild-type replicon containing the cassette (Northern blotting) and to determine if translation of the fusion protein occurred (GFP expression), (ii) RUBrep/GFP- $\Delta Not$ I to assay for *cis* complementation (both Northern gels and GFP expression), and (iii) RUBrep/CAT to assay for *trans* 

| Infectious cDNA    | <u>A construct</u> | Viability Co<br>P0 titer (plaque size) | <u>Trans</u><br>mplementation            |
|--------------------|--------------------|--|--|
|                    |                    |  | Construct +<br>RUBrep/GFP-ΔNotI<br>(GFP) |
| Robo502            |                    | 4.9 X 10 <sup>7</sup> (large)          | +  |
| Robo502-∆NotI      | -@·                | 1.7 X 10 <sup>7</sup> (small)          | +  |
| Robo502/IRES       |                    | 5.2 X 10⁵ (tiny)                       | +  |
| Robo502/IRES-ANotI | -@·                | 1.8 X 10⁵ (tiny)                       | +  |
| NS-C               | DRF IIII SP-ORF    | ☆ IRES                                 |  |

FIG. 2. Viability and *trans* complementation by infectious cDNA clone constructs. Shown are genomic diagrams of Robo502 and Robo502/ IRES, in which the junction-UTR between the NS- and SP-ORFs was replaced by the IRES of encephalomycarditis virus, without and with the *Not*I deletion in the P150 gene. As shown, all of these constructs give rise to viable virus, as shown by the titers and plaque morphologies of virus in transfection culture fluid harvested 4 days posttransfection. To test for the ability to complement the *Not*I deletion in *trans*, Vero cells were cotransfected with transcripts from one of these constructs and RUBrep/GFP- $\Delta Not$ I transcripts, and replication of RUBrep/GFP- $\Delta Not$ I was detected by GFP expression.

complementation (GFP expression in cells cotransfected with RUBrep/GFP- $\Delta NotI$ ). It was found that RUBrep/GFP containing all of these mutations replicated, and thus none adversely affected viability. When AUG1 or both AUG1 and AUG2 were mutated to AUA, neither translation of the fusion protein nor cis or trans complementation was detected; however, when AUG2 was mutated to AUA, translation and complementation occurred. These results indicate that translation of the C gene cannot initiate at AUG2. To use a different mutation to abrogate initiation of translation at the first AUG of the C gene, AUG1 was replaced with UAAUAA with the same result that translation of the fusion protein and complementation were not detected. However, when the UAAUAA was placed immediately upstream of AUG1, translation and complementation occurred. Next, 1- and 2-nt deletions were made in codon 6 of the C gene to maintain AUG1 in its wild-type context, including initiation of translation, but prevent normal translation of the C-GFP fusion protein. As expected, GFP expression was not detected in RUBrep/GFP, and neither mutant was capable of complementation. These results indicate that production of the C protein is necessary for complementation. Finally, to dissect the presence of C protein production from the effect of the RNA sequence at the 5' end of the C gene on complementation, silent mutations were made in each of the seven codons between AUG1 and AUG2; these mutations would preserve the amino acid sequence of the N terminus of the C protein but perturb the RNA sequence. None of the individual mutations had an effect on complementation, and when all of these mutations were combined in one mutant construct, complementation was still detected, providing further evidence that the C protein, rather than its encoding RNA, was the moiety responsible for complementation of the NotI deletion in P150.

### DISCUSSION

The goal of this study was to investigate whether a structural gene or protein of RUB could complement deletions of the *Not*I region of the NS-ORF that rendered replicons nonviable. The first indication of this phenomenon was a result in a

previous publication (32), in which we found that while the NotI deletion was lethal in replicons, virus with the same deletion was viable. Since the only difference between virus and replicons was the SP-ORF, we initiated this study with the hypothesis that one of the structural genes or proteins complemented the NotI gene or protein. By constructing  $\Delta NotI$ replicons that expressed via the SG RNA in-frame fusions of the individual structural proteins with GFP, we were able to determine that the moiety responsible for complementation of the NotI deletion was the C gene. A series of RUBrep/C-GFP fusions that contained progressive deletions of the C gene from either the 5' or 3' end, or both, showed that codons 1 to 277 of the C gene (in other words, the C gene lacking the E2 signal sequence) were required for replication. Similar results were obtained in both Vero and BHK cells, and thus complementation was not cell specific. We also found that the C gene of the related alphavirus SIN could not complement the NotI deletion, and thus complementation was specific to the RUB C gene. While initial results were based on GFP expression, Northern analysis revealed that in cells transfected with  $\Delta NotI$ replicon constructs that failed to express GFP, replicon-specific RNA was not detectable, indicating that complementation occurred at a basic step in the replication cycle, prior to accumulation of detectable replicon RNA, and not at the level of SG RNA synthesis and/or reporter gene translation.

As expected, since wild-type RUB was able to amplify replicons bearing *Not*I deletions (32), we also found that the C gene could complement RUBrep/GFP- $\Delta Not$ I in *trans* when expressed either from another replicon or from a plasmid, resulting in replication of RUBrep/GFP- $\Delta Not$ I. The complete C gene was not required for complementation in *trans*, since the 5' 31 codons of the C gene in RUBrep replicons or the 5' 58 codons of the C gene in a CMV-driven plasmid were sufficient for complementation. To some extent, if not completely, the difference between the amount of the C gene required for complementation in *cis* versus that required in *trans* was due to differences in efficiencies of virus production following transfection of the parent infectious clones. The RUBrep/C-GFP- $\Delta Not$ I constructs used to assay *cis* complementation were

| <u>Cassette</u>                                |   | <u>Translation/</u>                    | Comple                                       | ementation   |
|--|---|--|--|--|
|  |   | <b>Replication</b>                     | <u>Cis</u>                                   | <u>Trans</u>   |
|  |   | GFP cassette in<br>RUBrep<br>(GFP/RNA) | GFP cassette in<br>RUBrep-∆NotI<br>(GFP/RNA) | CAT cassette in<br>RUBrep+<br>RUBrep/GFP∆NotI<br>(GFP) |
| GFP/CAT  | -GFP/CAT  | +/+                                    | -/-  | _  |
| C-GFP/CAT                                      | -   | +/+                                    | +/+  | +  |
|  |   |  |  |  |
|  | AUGGCUUCUACUACCCCCAUCACC/   | AUG                                    |  |  |
| C(ΔAUG1)-GFP/CAT                               | AUÅGCUUCUACUACCCCCAUCACC/   | <u>AUG</u> —/+                         | -/-  |  |
| C(ΔAUG2)-GFP/CAT                               | AUGGCUUCUACUACCCCCAUCACC  | ₩* +/+                                 | +/+  | +  |
| C(AAUG1+2)-GFP/CAT                             | AUÅGCUUCUACUACCCCCAUCACC  | <u>wă</u> —/+                          | -/-  | -  |
| C(ZZAAUG1)-GFP/CAT                             | <u> </u>  | <u>vug</u> —/+                         | —/—  | <u></u>  |
| C(ZZ)-GFP/CAT                                  | ATCAUGGCUUCUACUACCCCCAUCACC   | AUG +/+                                | +/+  | +  |
| C(1-nt deletion)-GFP/CAT                       | AUGGCUUCUACUACCCCÅAUCACC  | <u>AUG</u> —/+                         | -/-  |  |
| C(2-nt deletion)-GFP/CAT                       | AUGGCUUCUACUACCCXXAUCACCA   | <u>.ug</u> —/+                         | -/-  |  |
| C(2 <sup>nd</sup> aa)-GFP/CAT                  | AUGGCAUCUACUACCCCCAUCACC  | <u>↓UG</u> +/+                         | +/+  | +  |
| C(3 <sup>rd</sup> aa)-GFP/CAT                  | AUGGCUUCĂACUACCCCCAUCACC  | <u>∿u</u> g <b>+/+</b>                 | +/+  | +  |
| C(3 <sup>rd</sup> aa)-GFP/CAT                  | <u>Αυσσου<sup>*</sup>δέλουλος του κατά τ</u> | <u>↓∪</u> g +/+                        | +/+  | +  |
| C(4 <sup>th</sup> aa)-GFP/CAT                  | AUGGCUUCUAC <sup>3</sup> ACCCCCAUCACC   | <u>↓∪</u> g <b>+/</b> +                | +/+  | +  |
| C(5 <sup>th</sup> aa)-GFP/CAT                  | AUGGCUUCUACUACÅCCCAUCACC  | <u>∧u</u> g <b>+/</b> ∔                | +/+  | +  |
| C(5 <sup>th</sup> aa)-GFP/CAT                  | AUGGCUUCUACUACĞCCCAUCACC/   | <u>AUG</u> +/+                         | +/+  | +  |
| C(6 <sup>th</sup> aa)-GFP/CAT                  | AUGGCUUCUACUACCCCAAUCACC  | <u>AUG</u> +/+                         | +/+  | +  |
| C(7 <sup>th</sup> aa)-GFP/CAT                  | AUGGEUUEUAEUAECEEEAUÅAECA   | <u>∿u</u> g <b>+/</b> +                | +/+  | +  |
| C(8 <sup>th</sup> aa)-GFP/CAT                  | AUGGEUUEUAEUAECEEEAUEAEÅ  | <u>↓∪</u> G +/+                        | +/+  | ÷  |
| C(2 <sup>nd</sup> -8 <sup>th</sup> aa)-GFP/CAT | <u>AUGGCÄUCÄACÄACÄCCÄAUÄACÄ</u>   | AUG +/+                                | +/+  | +  |

FIG. 3. Effect on complementation of mutations at the 5' end of the C gene. Using the C-GFP/CAT cassette, a series of mutations at the 5' end of the C gene were made. (The 5' end of the C gene has two in-frame AUGs separated by seven codons). Under the wild-type sequence, each mutated sequence is given with the mutation in boldface and designated with an asterisk. Each mutated cassette was assessed for replication in wild-type replicon RUBrep/GFP, *cis* complementation in RUBrep/GFP- $\Delta Not$ I, and *trans* complementation in RUBrep/CAT (in cells cotransfected with RUBrep/GFP- $\Delta Not$ I transcripts) as described in the legend to Fig. 1 and shown in Table 1.

based on Robo502, while the RUBrep/C-CAT constructs used to assay trans complementation were based on Robo503. Robo502 and Robo503 vary in the sequence of the additional nucleotides added to the transcript following the oligo(A) (A<sub>20</sub>) tract due to residual sequences in the restriction site used for linearization (EcoRI and SpeI, respectively): namely GAAUU in Robo502 transcripts and CUAG in Robo503 transcripts. While Robo502 and Robo503 transcripts have the same transfection efficiency, Robo503 transcripts replicate more rapidly than do Robo502 transcripts, which would hypothetically result in a more rapid and higher-level production of C protein. As predicted if this hypothesis were correct, when C-GFP fusions with 3' deletions were introduced into RUBrep/ CAT- $\Delta Not$ I, the construct with only the 5' 88 codons of the C gene replicated (in contrast to the 277 codons required for replication of RUBrep/C-GFP- $\Delta NotI$  constructs). Significantly, none of the constructs with 5' deletions of the C gene of any size were able to complement either in cis or in trans from any of the vectors employed.

While complementation is traditionally associated with proteins, RNA sequences or structures that function at long range in cis or function both in cis and in trans have been described for diverse RNA viruses (1a, 4, 5, 9, 11, 18, 24, 29, 31). Because the C genes in both Robo502/IRES and CMV-driven plasmids (neither of which synthesizes the RUB SG RNA nor contains the 5' UTR of the SG RNA) were found to be capable of complementing the NotI deletion, the SG RNA is not involved in complementation, and the complementation moiety resides within the coding sequences for the C protein. Concentrating on the 5' nine codons of the C gene, which deletion analysis indicated were essential for complementation, we found that mutations that abrogated translation of the C gene but only changed 1 nt within these codons also eliminated complementation, while mutations that changed the RNA sequence of these codons without changing the encoded amino acid sequence had no effect on complementation. Thus, the C protein and not the C gene RNA is the moiety responsible for complementation. Interestingly, in an alignment between the C proteins of RUB and SIN (Fig. 4A), the N terminus of the SIN protein is aligned with the second Met residue of the RUB protein, and the N-terminal eight residues of the RUB protein, which are essential for complementation, are not included in

| Α   |          |     |  |     |
|-----|----------|-----|--|-----|
| SIN | Capsid   | 1   | MNRGFFNMLGR-RPFPAPTAMWRPRRRRQAAPMPARNGLASQIQQLTTAVSALVIGQATRPQPPR  | 64  |
| RUB | Capsid   | 1   | MASTTPITMEDLQKALEAQSRALRAELAAGASQSRRPRPPRQRDSSTSGDDSGRDSGGPRRRRGNRGRGQRRDWSRAPPP                                       | 80  |
| SIN | Capsid   | 65  | PRPPPRQKKQAPKQPPK-PKKPKTQEKKKKQP-AKPKPGKRQRMALKLEADRLFDVKNEDGDVIGHALAM   | 132 |
| RUB | Capsid   | 81  | PEERQETRSQTPAPKPSRAPPQQPQPPRMQTGRGGSAPRPELGPPTNPFQAAVARGLRPPLHDPDTEAPTEACVTSWLWS * * * * * * * * * * * * * * * * * * * | 160 |
| SIN | Capsid   | 133 | EGKVMKPLHVKGTIDHPVLSKLKFTKSSAYDMEFAQLPVNMRSEAFTYTSEHPEGFYNWHHGAVQYSGGRFT   | 204 |
| RUB | Capsid   | 161 | EGEGAVFYRVDLHFTNLGTPPLDEDGRWDPALMYNPCGPEPPAHVVRAYNQPAGDVRGVWGKGERTYAEQDFRVGGTRWH                                       | 240 |
| SIN | Capsid   | 205 | IPRGVGGRGDSGRPIMDNSGRVVAIVLGGADEGTRTALSVVTWNSKGKTIKTTPEGTEEW 264   |     |
| RUB | Capsid   | 241 | RLLRMPVRGLDGDSAPLPPHTTERIETRSARHPWRIRFGAPQAFLAGLLLATVAVGTARA 300   |     |
| R   |          |     |  |     |
| BUB | Capsid   | 1   | MASTTPITMEDLOKALEAOSRALRAELAAGASOSRRPRPP   | 40  |
| Not | T Domain | 1   |  | 80  |
| NOC | i Domain | 1   |  | 00  |
| RUB | Capsid   | 41  | RQRDS-STSGDDSGRDSGGPRRRRGN   | 88  |
| Not | I Domain | 81  | PERPWADGGARAWAKFFRGCAWAQRLLGEPAVMHLPYTDGDVPQLIALALRTLAQQGAALALSVRDLPGGAAFDANAVTA 1                                     | 60  |

.\* \* RUB Capsid 89 SQTPAPKPSRAP 100 NotI Domain161 AVRAGPRQSAAA 172 \*. \* \*

FIG. 4. Alignment of the sequences of RUB C protein with those of the SIN C protein and *Not*I domain. The entire RUB and SIN C protein sequences are included in the alignment; amino acids are numbered from the N-terminal Met residue. In the RUB C X *Not*I alignment, only the N-terminal 100 residues of the C protein are used. The numbering of the *Not*I domain represents the residue number within the P150 protein. The alignment was made with the MacVector version 6.5.3 program.

the alignment. (Overall, in this alignment, the two C proteins share 17% identity and 26% similarity.)

The mechanism by which the C protein complements the NotI region of the P150 replicase protein was not investigated in this study; however, complementation occurred at an early step in the replication cycle, since RUBrep/GFP- $\Delta Not$ I failed to accumulate detectable replicon-specific RNA. Since complementation occurred in cis, initial replication of input replicon transcripts would theoretically have to occur to allow for synthesis of the C protein (none was introduced by infecting virions, since this study was done entirely with transcripts), although translation of C protein from partially degraded transcripts cannot be completely ruled out. If initial replication is required to produce C protein, then complementation operates in the process of amplification, rather than initiation, of virus RNA synthesis. Interestingly, it has been reported that the C protein colocalized with P150 on tubular structures in RUB-infected cells late in infection (13). The C protein contains a motif between residues 28 and 56 that binds the RUB genomic RNA (21), and RNA binding by the C protein has been shown to be regulated by a phosphorylation site at Ser-46 (14). Whether these activities assist in RNA replication as well as playing a role in their obvious function of encapsidation is not known; it is conceivable that C protein binding of the genome RNA is a factor in efficient release from the replication complex. However, since only the N-terminal 31 residues of the C gene were needed for complementation in trans by RUBrep/C-CAT, these activities may not be involved in complementation of the NotI deletion. Besides a specific step

in the viral replication cycle, C protein complementation of the *Not*I region of P150 may involve binding to cellular factors. The C protein has been shown to interact with two mitochondrial proteins, p32 and Par-4, and the p32-binding domain within the C protein has been mapped to the N-terminal region of the protein (2, 23). While the function and utility of these interactions in RUB replication have not been elucidated, p32 belongs to the family of cellular defense collagens (30). The C protein has also been associated with induction of apoptosis in RK13 cells, a cell line exquisitely sensitive to RUB-induced cytopathic effect (6).

Finally, it is possible that C protein complementation of the NotI domain of the P150 protein is mediated through some general activity rather than direct compensation for NotI domain function: for example, binding to transcripts and stabilizing them or efficiently targeting them. This would be a function that the C protein in virions possibly plays during virus infection, and under this scenario, replicons with NotI deletion would simply replicate too inefficiently to establish themselves before degradation occurred or some cellular defense mechanism was fully induced. In this regard, the function of the NotI domain is not known (since it has no detectable homology with replicase proteins of other viruses), and it is notable that when aligned with the NotI domain, the N terminus of the C protein shares little homology (22% similarity and 16% identity) with this domain (Fig. 4B), and we found that neither P150 nor the NotI domain could complement  $\Delta NotI$  replicons when expressed in trans.

### ACKNOWLEDGMENTS

Support for this research was provided by a grant from the National Institutes of Health (AI21389).

We thank Ping Jiang for sequencing and Kai-Hui Wu for making alignments.

#### REFERENCES

- Adams, S. D., W.-P. Tzeng, M.-H. Chen, and T. K. Frey. 2003. Analysis of intermolecular RNA-RNA recombination by rubella virus. Virology 309: 258–271.
- 1a.Barry, J. K., and W. A. Miller. 2002. A -1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. Proc. Natl. Acad. Sci. USA 99:11133-11138.
- Beatch, M. D., and T. C. Hobman. 2000. Rubella virus capsid associates with host cell protein p32 and localizes to mitochondria. J. Virol. 74:5569–5576.
- Chen, J.-P., J. H. Strauss, E. G. Strauss, and T. K. Frey. 1996. Characterization of the rubella virus nonstructural protease domain and its cleavage site. J. Virol. 70:4707–4713.
- Choi, I. R., and K. A. White. 2002. An RNA activator of subgenomic mRNA1 transcription in tomato bushy stunt virus. J. Biol. Chem. 277:3760–3766.
- Choi, I. R., M. Ostrovsky, G. Zhang, and K. A. White. 2001. Regulatory activity of distal and core RNA elements in Tombusvirus subgenomic mRNA2 transcription. J. Biol. Chem. 276:41761–41768.
- Duncan, R., A. Esmaili, L. M. Law, S. Bertholet, C. Hough, T. C. Hobman, and H. L. Nakhasi. 2000. Rubella virus capsid protein induces apoptosis in transfected RK13 cells. Virology 275:20–29.
- Frey, T. K. 1994. Molecular biology of rubella virus. Adv. Virus Res. 44:69– 160.
- Gros, C., and G. Wengler. 1996. Identification of an RNA-stimulated NT-Pase in the predicted helicase sequence of the Rubella virus nonstructural polyprotein. Virology 217:367–372.
- Guo, L., E. M. Allen, and W. A. Miller. 2001. Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. Mol. Cell 7:1103–1109.
- Hahn, C. S., Y. S. Hahn, T. J. Braciale, and C. M. Rice. 1992. Infectious Sindbis virus transient expression vectors for studying antigen processing and presentation. Proc. Natl. Acad. Sci. USA 89:2679–2683.
- Kim, K. H., and C. L. Hemenway. 1999. Long-distance RNA-RNA interactions and conserved sequence elements affect potato virus X plus-strand RNA accumulation. RNA 5:636–645.
- Koonin, E. V., A. E. Gorbalenya, M. A. Purdy, M. N. Rozanov, G. R. Reyes, and D. W. Bradley. 1992. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. Proc. Natl. Acad. Sci. USA 89:8259–8263.
- Kujala, P., T. Ahola, N. Ehsani, P. Auvinen, H. Vihinen, and L. Kääriäinen. 1999. Intracellular distribution of rubella virus nonstructural protein P150. J. Virol. 73:7805–7811.
- Law, L. M., J. C. Everitt, M. D. Beatch, C. F. Holmes, and T. C. Hobman. 2003. Phosphorylation of rubella virus capsid regulates its RNA binding activity and virus replication. J. Virol. 77:1764–1771.
- Liang, Y., J. Yao, and S. Gillam. 2000. Rubella virus nonstructural protein protease domains involved in *trans-* and *cis-cleavage activities*. J. Virol. 74:5412–5423.

- Liang, Y., and S. Gillam. 2000. Mutational analysis of the rubella virus nonstructural polyprotein and its cleavage products in virus replication and RNA synthesis. J. Virol. 74:5133–5141.
- Liang, Y., and S. Gillam. 2001. Rubella virus RNA replication is *cis* preferential and synthesis of negative- and positive-strand RNAs is regulated by the processing of nonstructural protein. Virology 282:307–319.
- Lindenbach, B. D., J.-Y. Sgro, and P. Ahlquist. 2002. Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. J. Virol. 76:3905–3919.
- Liu, X., S. L. Ropp, R. J. Jackson, and T. K. Frey. 1998. The rubella virus nonstructural protease requires divalent cations for activity and functions in *trans.* J. Virol. 72:4463–4466.
- Liu, X., J. Yang, A. M. Ghazi, and T. K. Frey. 2000. Characterization of the zinc binding activity of the rubella virus nonstructural protease. J. Virol. 74:5949–5956.
- Liu, Z., D. Yang, Z. Qiu, K.-T. Lim, P. Chong, and S. Gillam. 1996. Identification of domains in rubella virus genomic RNA and capsid protein necessary for specific interaction. J. Virol. 70:2184–2190.
- Marr, L. D., C.-Y. Wang, and T. K. Frey. 1994. Expression of the rubella virus nonstructural protein ORF and demonstration of proteolytic processing. Virology 198:586–592.
- Mohan, K. V., B. Ghebrehiwet, and C. D. Atreya. 2002. The N-terminal conserved domain of rubella virus capsid interacts with the C-terminal region of cellular p32 and overexpression of p32 enhances the viral infectivity. Virus Res. 85:151–161.
- Paul, A. V., J. Peters, J. Mugavero, J. Yin, J. H. van Boom, and E. Wimmer. 2003. Biochemical and genetic studies of the VPg uridylylation reaction catalyzed by the RNA polymerase of poliovirus. J. Virol. 77:891–904.
- Pugachev, K. V., M. S. Galinski, and T. K. Frey. 2000. Infectious cDNA clone of the RA27/3 vaccine strain of Rubella virus. Virology 273:189–197.
- Pugachev, K. V., W.-P. Tzeng, and T. K. Frey. 2000. Development of a rubella virus vaccine expression vector: use of a picornavirus internal ribosome entry site increases stability of expression. J. Virol. 74:10811–10815.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Seed, B., and J. Y. Sheen. 1988. A simple phase-extraction assay for chloramphenicol acyltransferase activity. Gene 67:271–277.
- Sit, T. L., A. A. Vaewhongs, and S. A. Lommel. 1998. RNA-mediated transactivation of transcription from a viral RNA. Science 281:829–832.
- Tenner, A. J. 1999. Membrane receptors for soluble defense collagens. Curr. Opin. Immunol. 11:34–41.
- Tiley, L., A. M. King, and G. J. Belsham. 2003. The foot-and-mouth disease virus *cis*-acting replication element (*cre*) can be complemented in *trans* within infected cells. J. Virol. 77:2243–2246.
- Tzeng, W.-P., M.-H. Chen, C. A. Derdeyn, and T. K. Frey. 2001. Rubella virus DI RNAs and replicons: requirement for nonstructural proteins acting in *cis* for amplification by helper virus. Virology 289:63–73.
- Tzeng, W.-P., and T. K. Frey. 2002. Mapping the rubella virus subgenomic promoter. J. Virol. 76:3189–3201.
- 34. Wang, X., Y. Liang, and S. Gillam. 2002. Rescue of rubella virus replicationdefective mutants using vaccinia virus recombinant expressing rubella virus nonstructural proteins. Virus Res. 86:111–122.
- Yao, J., D. Yang, P. Chong, D. Hwang, Y. Liang, and S. Gillam. 1998. Proteolytic processing of rubella virus nonstructural proteins. Virology 246: 74–82.