JOURNAL OF VIROLOGY, Aug. 1991, p. 4017–4025
0022-538X/91/084017-09$02.00/0

Vol. 65, No. 8

Recombination between Sindbis Virus RNAs

BARBARA G. WEISS AND SONDRA SCHLESINGER*

Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110-1093

Received 14 February 1991/Accepted 16 April 1991

The genome (49S RNA) of Sindbis virus is a positive-strand RNA of 11.7 kb that consists of two domains. The 5′ two-thirds of the RNA codes for the proteins required for replication and transcription of the RNA. The 3′ one-third codes for the structural proteins. The latter are translated from a 26S subgenomic RNA identical in sequence to the 3′ one-third of the genome. The 26S RNA is transcribed by initiation from an internal promoter that spans the junction between the nonstructural and structural genes. We have used Sindbis virus RNAs transcribed from cloned cDNAs to demonstrate recombination between Sindbis virus RNAs in cultured cells. Several different combinations of deleted or mutationally altered RNAs gave rise to infectious recombinants. In 7 of 10 different crosses, the infectious recombinant RNAs were larger than wild-type 49S RNA. We sequenced the recombinant RNAs in the region spanning the junction between the nonstructural and structural protein genes from five different crosses. In three of the crosses, this is the only region within which recombination could have taken place to produce an infectious 49S RNA. Recombination also occurred in this region in the other two crosses. The recombinant RNAs were distinct from wild-type RNA and from each other. All contained sequence insertions derived from the parental RNAs. One contained a deletion and a rearrangement, and one also contained a stretch of 11 nucleotides not found in the Sindbis virus genome. When each of the parental RNAs contained a functional subgenomic RNA promoter, both promoters were present and functional in the recombinant RNA. Those recombinants with large sequence insertions showed evidence of evolution toward the wild-type single-junction RNA.

The exchange of genetic information among RNA viruses can occur by two different mechanisms. One mechanism, the reassortment of genes, is limited to those viruses with segmented genomes. The other, recombination between RNA molecules, was reported almost 30 years ago with mutants of poliovirus (12). More recently, this type of recombination has also been described for aphthoviruses (16), coronaviruses (18), and the plant bromoviruses brome mosaic virus (4) and cowpea chlorotic mottle virus (2). In contrast to these results, initial attempts to detect recombination between temperature-sensitive mutants of the alphavirus Sindbis virus in cultured cells were unsuccessful (5). In those experiments, the relatively high rate of reversion of the temperature-sensitive mutants could have obscured a low frequency of recombination. The possibility of identifying recombinants between Sindbis virus RNAs in the absence of any revertants became feasible with the development of cloned Sindbis virus cDNAs that can be transcribed into biologically active RNAs (21, 28).

The genome (49S RNA) of Sindbis virus is a single strand of RNA of positive polarity. It consists of 11.7 kb plus a poly(A) tail (30, 31). The 5′ two-thirds of the RNA codes for the proteins required for replication and transcription of the RNA. The 3′ one-third codes for the structural proteins: the capsid protein and the proteins that comprise the envelope of the virion. The nonstructural proteins are translated from genomic-length mRNAs; the structural proteins are translated from a subgenomic RNA (26S RNA) identical in sequence to the 3′ one-third of the genome. This subgenomic RNA is transcribed from the minus strand of genomic RNA by initiation from an internal promoter that spans the junction between the structural and nonstructural genes (20, 26).

We describe here recombination between Sindbis virus RNAs, using parental RNAs that have been transcribed from engineered cDNAs. One of the parental RNAs (TRCAT) contains only the nonstructural genes and the *cis*-acting sequences of the Sindbis virus genome. The structural genes have been replaced by the bacterial gene encoding chloramphenicol acetyltransferase (CAT) (35; Fig. 1). When this RNA is transfected into cells, both genomic and subgenomic RNAs are produced, but the genomic RNA is not packaged.

We recently reported that when TRCAT and a defective interfering (DI) RNA able to produce a subgenomic RNA coding for the viral structural proteins are cotransfected into cells, the two RNAs complement each other to form infectious virions which contain a segmented genome (7). In some instances, we also observed an RNA similar in size to the 49S genomic RNA, suggesting that recombination between TRCAT and the DI RNA occurred. We have now investigated recombination between Sindbis virus RNAs in more detail not only with these RNAs but also with other deleted and mutated genomic RNAs. We sequenced recombinant RNAs in the region spanning the junction between the nonstructural and structural protein genes from five different crosses. In three of the crosses, this is the only region within which recombination could have taken place to produce an infectious 49S RNA. Recombination also occurred in this region in the other two crosses. The recombinant RNAs were distinct from wild-type RNA and from each other. All contained sequence insertions derived from the parental RNAs. One contained a deletion and a rearrangement, and one also contained a stretch of 11 nucleotides not found in the Sindbis virus genome. When each of the parental RNAs contained a functional subgenomic RNA promoter, both promoters were present and functional in the recombinant RNA.

* Corresponding author.

MATERIALS AND METHODS

Plasmid constructions. DI cDNAs containing the structural protein genes of 26S mRNA were constructed from either KDI25 (21) or CTS14 (34) and the 49S cDNA clone Toto1102 (28). All DNA fragments were purified by agarose gel electrophoresis before ligations.

DI(26S) and DI(26S.1). DI(26S) (diagrammed in Fig. 6) has been described previously (7). DI(26S.1) (diagrammed in Fig. 1) was constructed from CTS14 and Toto1102. CTS14 was cut at the *Eco*RI site at position 931, filled in with Klenow polymerase, and then cut with *Xho*I at position 2716. The resulting 2,827-bp fragment was ligated to a 4,250-bp *Ssp*I-*Xho*I fragment from Toto1102. The latter fragment included 102 nucleotides upstream of the start of 26S RNA and the complete sequence of 26S mRNA. The promoter for the SP6 polymerase is located directly upstream of the start of the DI RNA sequence (21). The DNA, linearized with *Xho*I, was transcribed into an RNA of 5,186 nucleotides. The only significant difference between the two DI cDNAs with respect to this study was that CTS14 contained sequences from the bacterial gene encoding CAT and was used because of convenient restriction sites.

DI26S.1-36 and DI26S.1-39. Both clones, derived from CTS14 and Toto1102, were constructed so that the DI RNA contained the coding sequences from 26S RNA but lacked the promoter. CTS14 was linearized at an *Nco*I site at position 1232, blunt ended with mung bean nuclease, and then cut with *Xho*I. The 3,285-bp fragment was ligated to an *Xba*I-*Xho*I fragment derived from Toto1102 (from nucleotides 7612 to 11749) in which the *Xba*I site had been blunt ended with mung bean nuclease. Sequence analysis across the *Nco*I-*Xba*I ligation site for DI26S.1-36 showed that nucleotides from 1 to 1233 were derived from CTS14 and were linked to nucleotide 7616 in Toto1102. The region surrounding the junction was deleted, but the last two 3′ nucleotides of the *Xba*I linker present in Toto1102 were retained. The 26S RNA sequences present in this clone begin at position +19 in Toto1102 26S RNA. The sequence of DI26S.1-39 showed that the first 1,233 nucleotides of CTS14 were linked to nucleotide 7628 in Toto1102. The 26S sequences present in this clone begin at position +30 in the Toto1102 26S RNA.

Toto plasmids containing deletions in the nonstructural proteins. The clones were constructed from Toto1101 (28) by C. M. Rice and his colleagues. CR1.8 has an in-frame deletion extending from nucleotides 425 to 574 in the nsP1 (ns indicates nonstructural) gene; CR2.4 has an in-frame deletion from nucleotides 2130 to 2201 in the nsP2 gene; CR3.4 has an in-frame deletion from nucleotides 4755 to 4919 in the nsP3 gene; CR4.10 has an in-frame deletion from nucleotides 6878 to 7031 in the nsP4 gene.

Toto E2C$_{415}$C$_{416}$. The clone was constructed by K. G. Nitschko and M. Schlesinger from Toto1000. Site-specific mutations in the E2 protein gene were made to convert the cysteine at position 415 to serine and the cysteine at position 416 to alanine. RNA transcribed from this cDNA did not form infectious virions.

TRCAT and TSCATts6. The clones have been described previously (35). They contain the CAT gene replacing the Sindbis virus structural protein genes between nucleotides 7612 and 11087 in Toto1002 (28) for TRCAT (diagrammed in Fig. 1) and nucleotides 7715 to 11087 in Toto1000 for TSCATts6.

Transcription, transfection, and passaging of virus. Transcriptions were carried out with the SP6 DNA-dependent RNA polymerase as directed by the supplier. All transcripts were labeled with [^{3}H]UTP and capped during transcription.

RNA transfections were performed essentially as described previously (34) except that we used 35-mm instead of 60-mm culture dishes. Monolayers of chicken embryo fibroblasts (CEF) were transfected with 0.5 μg of each transcript and lipofectin (7.5 μg) in a volume of 0.2 ml of Eagle minimal essential medium (MEM). After 1 h at 30°C, 0.8 ml of serum-free MEM containing dactinomycin (1 μg/ml) was added. One hour later, the 1 ml was removed and replaced with 1.5 ml of MEM containing 3% fetal calf serum, 1 μg of dactinomycin per ml, and 33 μC of [^{3}H]uridine per ml. After overnight incubation at 30°C, the extracellular fluid was harvested and the cellular RNA was isolated. Transfections were designated as passage 1. Passage 2 was obtained by infection of new monolayers of CEF with a sample of the passage 1 stock (usually 0.2 ml). The infected cells were labeled as described above.

Purification of virus from plaques. Samples of the extracellular fluids were titered on monolayers of CEF with an overlayer of MEM and 1% fetal calf serum in 0.75% agarose. Infectious virus was obtained from individual plaques by cutting out the agarose piece containing the plaque and eluting the virus into 0.5 ml of MEM containing 3% fetal calf serum.

cDNA synthesis, PCR amplification, and cloning and sequencing of the region surrounding the junction domains. The methods used were modifications of those described by Grakoui et al. (9). The Moloney murine leukemia virus reverse transcriptase (200 U per reaction) was used for cDNA synthesis. RNA isolated from CEF infected with a sample of virus that had been isolated from a single plaque was used as a template for first-strand synthesis. Reactions were carried out at 37°C for 20 to 60 min in the presence of RNAsin (400 U/ml). For polymerase chain reactions (PCR) involving the synthesis of products in the range of 200 to 300 bp, the cDNA and RNA products were first heated at 90°C for 2 min and then the reaction volume was doubled by adding 10 μl of a solution containing 1× PCR buffer (50 mM KCl, 10 mM Tris [pH 8.3], 2.5 mM $MgCl_2$, 0.1% gelatin), 20 pmol each of positive- and negative-polarity primers, 0.5 mM each of the four deoxynucleotides, and 2.5 U of *Taq* DNA polymerase. A drop of mineral oil was added to each mixture, and a 30-cycle PCR was carried out. Each cycle involved a 1-min denaturation at 94°C, followed by 1 min of annealing at 41 to 72°C, depending on the primer pair used, and 3 min of polymerization at 72°C. The reaction products were blunt ended with T4 DNA polymerase and phosphorylated with T4 polynucleotide kinase. Depending on size, the product was purified by electrophoresis through either a nondenaturing 6% polyacrylamide gel or a 0.8% agarose gel. After staining with ethidium bromide, the major product was cut out of the gel and isolated by electrophoresis onto DEAE paper. Products separated in polyacrylamide gels were cut out of the gel and electrophoresed onto DEAE paper after insertion into a slot of an agarose gel. DNA was eluted from the paper with 1 M NaCl–0.1 mM EDTA–20 mM Tris HCl (pH 7.5), precipitated in the presence of 5 μg of carrier tRNA, and cloned into the *Sma*I site of M13mp19. Recombinant phage plaques containing insertions of the correct size were sequenced by the dideoxy-chain termination method (29) using a Sequenase (version 2.0) kit from United States Biochemical Corp. and either Sindbis virus-specific primers or a 17-nucleotide universal primer derived from nucleotides 6292 to 6308 in M13mp19.

When the PCR product was expected to be larger than 700

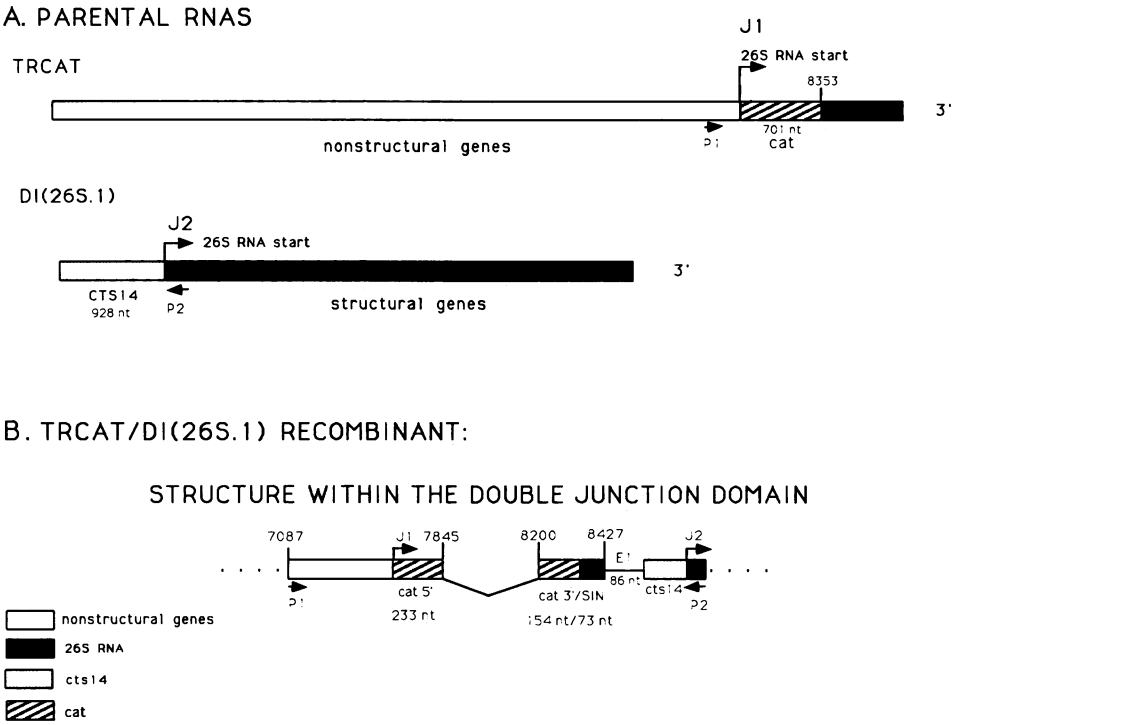

FIG. 1. Diagrams of the parental RNAs, TRCAT and DI(26S.1), and of the double-junction recombinant, TRCAT/26S.1, in the region of the double-junction domain. (A) Parental RNAs, TRCAT and DI(26S.1). J1 refers to region spanning the subgenomic RNA promoter in TRCAT; J2 refers to this region in DI(26S.1). Primer 1 (P1) is a positive-sense, 21-nucleotide (nt) oligomer identical to nucleotides 7087 to 7107 in the 3′ domain of the nsP4 gene. This sequence is not present in DI(26S.1) or 26S RNA. Primer 2 (P2) is a negative-sense oligomer containing 16 nucleotides complementary to nucleotides 7632 to 7646 of Toto1100 located at the 5′ end of the capsid gene. These nucleotides are not present in the negative strand of TRCAT. (B) The double-junction recombinant, TRCAT/26S.1, in the region of the double-junction domain. J1 and J2 are the same regions as those described for TRCAT and DI(26S.1), respectively. The 86-nucleotide segment from the E1 gene inserted into the recombinant RNA is indicated by a thin line between the two parental RNA domains.

bp, the RNA present in the reverse transcriptase reactions was hydrolyzed by treatment with alkali, and the cDNA was precipitated with ethanol before PCR amplification (8). The PCR reaction mixture contained 50 pmol of each primer, the four deoxyribonucleotides each at a concentration of 0.2 mM, and 1× PCR buffer (specified above) in a final volume of 50 μl. The sample was heated for 1 min at 100°C before addition of enzyme, and the first-cycle extension was carried out at 72°C for 40 min (6). These conditions markedly and reproducibly enhanced the amplification of large cDNA products. More specific amplification was also obtained by using, during reverse transcription, a minus-strand primer different from the one used in the PCR reaction.

RESULTS

Recombination between TRCAT and DI genomes. The structures of TRCAT and a DI RNA [DI(26S.1)] containing the viral structural genes placed downstream of the promoter for the transcription of the subgenomic RNA are diagrammed in Fig. 1A. When TRCAT and DI(26S.1) were transfected into CEF, both RNAs and their respective subgenomic RNAs were synthesized (Fig. 2A, lane 1). This result would be expected if some of the cells were transfected with both RNAs, since TRCAT provides the proteins required for the replication and transcription of both RNAs (7). When cells were transfected with TRCAT alone, both TRCAT genomic RNA and its subgenomic RNA were synthesized. No virus-specific RNAs were synthesized in cells transfected with DI(26S.1) RNA alone (data not shown).

The cotransfected cells, but not cells transfected with either of the RNAs alone, gave rise to titers of 2×10^5 to 8×10^5 PFU/ml. These plaques were the result of the ability of the two RNAs to function together as a segmented genome and to be copackaged (7). TRCAT and DI(26S.1) RNAs were both packaged (Fig. 2A, lane 2). A sample of the extracellular fluid from the transfected cells (passage 1) was used to infect a new monolayer of CEF, and the virus-specific RNAs synthesized in the infected cells were analyzed by agarose gel electrophoresis. Both TRCAT and DI(26S.1) as well as their subgenomic RNAs were synthesized. In addition to these RNAs, an RNA similar in size to 49S virion RNA was also detected and was still present after another passage (lane 3), indicating that it had also been packaged, but it was not enriched over the other viral genomic RNAs.

We thought that this might be a recombinant 49S genomic RNA, but we could not enrich for particles containing the putative recombinants by isolating virus particles from independent plaques. Instead, particles purified from plaques were ones containing the segmented genome (data not shown, but see reference 7). The strategy that we used to obtain the 49S-like RNA (Fig. 2A, lane 2) free of TRCAT and DI RNAs was to subject the RNAs, without prior denaturation, to agarose gel electrophoresis, elute the 49S-like RNA from the gel, and then transfect it into CEF to obtain a stock of infectious particles. These particles were then used to

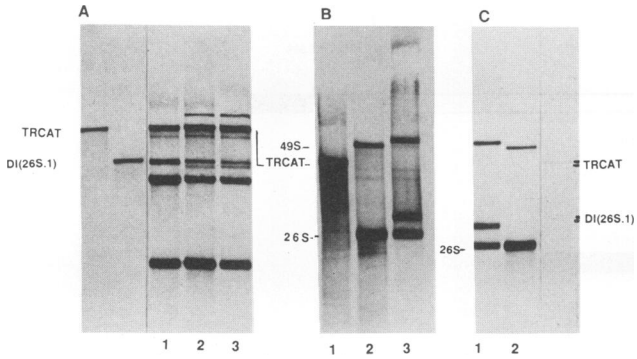

FIG. 2. Complementation and recombination following cotransfection of cells with TRCAT and DI(26S.1) RNAs. (A) Patterns of Sindbis virus RNAs synthesized in transfected and infected cells. Lanes: 1, pattern of [^{3}H]uridine-labeled RNAs synthesized in monolayers of CEF transfected with TRCAT and DI(26S.1); 2 and 3, patterns of RNAs synthesized during the formation of passages 2 and 3, respectively. The two left lanes show the position of migration of the parental RNA transcripts, TRCAT and DI(26S.1). The smallest RNA bands in lanes 1 to 3 are the subgenomic RNAs transcribed from the negative strand of TRCAT. (B) Patterns of viral RNA synthesized in cells infected with virus obtained from gel-purified RNA. The intracellular 49S-like RNA (see lane 2, panel A), was isolated from an agarose gel, without prior denaturation, and transfected into CEF. The virus released from the transfected cells was used to infect a new monolayer of CEF, and [^{3}H]uridine-labeled Sindbis virus-specific RNAs are seen in lane 3. Lane 1 shows the position of migration of TRCAT; lane 2 shows the positions of migration of authentic 49S and 26S RNAs. (C) Patterns of viral RNAs in cells infected with plaque-purified recombinant virus. A sample of the virus used to obtain the RNA pattern seen in panel B, lane 3, was titered, and virus stocks were prepared from independent plaques. The RNA patterns seen in lanes 1 and 2 represent infections by two different plaque isolates. The third lane (not labeled) shows the migration of the parental RNA transcripts.

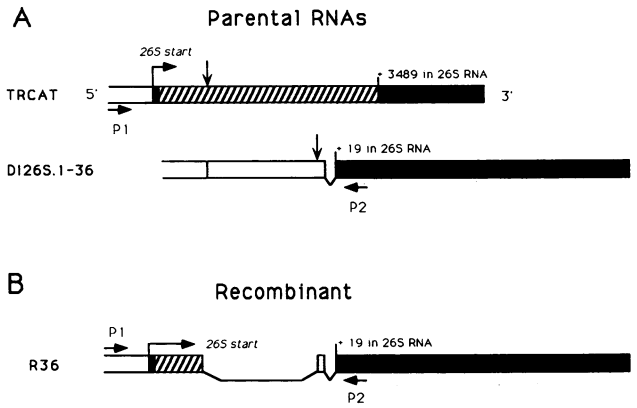

FIG. 3. Diagram of a cross between TRCAT and DI26S.1-36. (A) Parental RNAs; (B) recombinant. The cDNA from which this DI RNA was derived is described in Materials and Methods. Primer 1 (P1) is a positive-sense, 21-nucleotide oligomer identical to nucleotides 7463 to 7483 in the 3'-terminal region of nsP4. Primer 2 (P2) is a negative-sense oligomer of 20 nucleotides that are complementary to nucleotides 7690 to 7709 located at the 5' end of the capsid gene. Vertical arrows indicate the breakpoint in each of the parental RNAs that would give rise to the recombinant structure shown. Symbols: ▦, CAT sequences in DI; ▨, CAT sequences in TRCAT; ■, 26S RNA sequence.

infect a new monolayer of cells, and the resulting RNA pattern is shown in Fig. 2B, lane 3. Neither TRCAT RNA nor DI(26S.1) RNA was present, but there was a significant amount of an RNA with an estimated size of 5 kb migrating slower than the 26S RNA. Furthermore, the 49S-like RNA was larger than authentic 49S RNA (Fig. 2B, lane 2).

We originally thought that the 5-kb RNA might be a DI RNA and, if so, should be lost upon plaque purification of the infectious particles. Three types of particles were obtained from plaques: those that gave the same pattern of viral RNAs in infected cells as had the original virus particles (Fig. 2C, lane 1), those that produced only RNAs identical in size to authentic Sindbis virus 49S and 26S RNAs (Fig. 2C, lane 2), and those that appeared to be a mixed population containing both of the genomic RNAs seen in lanes 1 and 2 as well as the 5-kb and 26S RNAs (data not shown, but see Fig. 4 for examples of mixed populations). There was a direct correlation between the presence of the RNA larger in size than 49S RNA and the 5-kb RNA. This led us to suspect that this genomic RNA contained two promoters for subgenomic RNA transcription and that the 5-kb RNA was a second subgenomic RNA. On the basis of this assumption, we generated PCR products by using the primers indicated in Fig. 1. The PCR product obtained from authentic 49S RNA was 569 bp in length; the product obtained from this larger 49S-like RNA was 1.4×10^3 bp in length. Sequence analysis of this cDNA established that the subgenomic RNA promoter was present in duplicate (diagrammed in Fig. 1B). The upstream sequences (from nucleotides 7087 to 8427) were derived from TRCAT, but there was a 354-nucleotide in-frame deletion within the CAT gene. These sequences were followed by an 86-nucleotide insertion derived from the E1 structural protein gene. This inserted sequence, derived from nucleotides 4517 to 4603 of 26S RNA, was present in DI(26S.1) but not in TRCAT. The sequences downstream of the E1 insertion were derived from DI(26S.1) beginning at nucleotide 744 from the 5' terminus. The nucleotide sequence of the recombinant showed that multiple events had occurred in the generation of this RNA.

We also constructed two DI cDNAs that contained the entire coding region of 26S RNA but lacked the promoter for its transcription. RNAs transcribed from either of these cDNAs and transfected into cells in conjunction with TRCAT RNA were replicated, but transcription of 26S subgenomic RNA should not occur in the absence of recombination. The recovery of PFU from cotransfected cells suggested that recombination had occurred, and this was verified by further analysis. Cells infected with several independent plaque isolates synthesized only 49S and 26S viral RNAs. The recombinant virus from the cross between TRCAT and DI26S.1-36 synthesized a 26S RNA that migrated with a mobility slower than that of wild-type 26S RNA. One plaque-purified recombinant from each cross was subjected to PCR analysis. The major DNA product obtained by amplification was cloned into M13mp19, and one clone from each cross was sequenced. These recombinant RNAs had only a single promoter for subgenomic RNA synthesis, but they both retained sequences from the CAT gene (Fig. 3 shows a diagram of one of these crosses). Both recombinant RNAs could have been generated by a single crossover event in the CAT gene. In both examples, there was no homology at the crossover site.

Recombination between TRCAT and Sindbis virus genomic RNAs deleted in the nonstructural genes. The four proteins coded by the nonstructural protein genes of Sindbis virus are translated as a polyprotein that is co- and posttransla-

TABLE 1. Crosses between Sindbis virus RNAs that gave rise to infectious recombinants

Parental RNAs[a]	Size of 49S RNAs[b]	Two subgenomic RNAs observed
TRCAT × DI(26S.1)	>wt	+ (Fig. 1 and 2)
TRCAT × DI26S.1-36	>wt	− (Fig. 3)
TRCAT × DI26S.1-39	>wt	−
TSCATts6 × CR1.8	wt	−
TRCAT × CR1.8	>wt	+
TSCATts6 × CR2.4	wt	−
TRCAT × CR3.4	>wt	+ (Fig. 4 and 5)
TSCATts6 × CR4.10	wt	−
DI(26S) × $E2C_{415}C_{416}$	>wt	+ (Fig. 6 and 7)
$E2C_{415}C_{416}$ × CR2.4	>wt	+

[a] CR1.8, CR2.4, CR3.4, and CR4.10 are in-frame deletion mutations in the nsP1, nsP2, nsP3, and nsP4 genes, respectively of Sindbis virus Toto1101 RNAs. They and the other parental RNAs are described in Materials and Methods.

[b] >wt is defined as having a mobility slower than that of authentic 49S RNA in agarose gel electrophoresis except for the recombinants between TRCAT and DI26S.1-36 or DI26S.1-39. The latter two RNAs migrated with 49S RNA in agarose gels but were shown by sequence analysis to have inserts in their genomes.

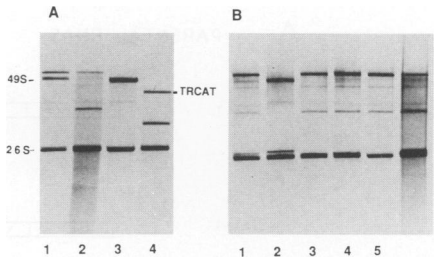

FIG. 4. Recombinant RNA profiles from a cross between TRCAT and CR3.4. (A) Patterns of viral RNAs in cells infected with virus stocks obtained from independent plaques. These plaque isolates were obtained by titration of the extracellular fluid from cells cotransfected with TRCAT and CR3.4. Lanes: 1 and 2, RNA patterns for two independent plaque isolates; 3, RNA pattern from cells infected with wild-type Sindbis virus; 4, positions of migration of TRCAT and DI(26S) RNAs included as reference markers. (B) Assay in which the plaque-purified virus shown in panel A, lane 2, was titered and virus eluted from several independent plaques was used to infect CEF. Lanes 1 to 5 show the RNA patterns observed for five independent plaques. The right lane shows the RNA pattern for cells infected with the original virus population and is identical to that displayed in panel A, lane 2.

tionally cleaved by an autoprotease (31). In-frame deletions in each of the genes were constructed in the viral cDNA (27a). RNAs transcribed from these DNAs do not produce virus. We tested four of the RNAs, each with a deletion in one of the nonstructural protein genes, for their ability to recombine with TRCAT or the closely related TSCATts6. TSCATts6 was originally chosen as a parental RNA so that we could use the RNA^- temperature-sensitive phenotype as a marker in recombination. Because many of the recombinant RNAs contained inserts which might contribute to a temperature-sensitive phenotype, we did not evaluate the distribution of this phenotype among the recombinants.

For each cross, the transcribed RNAs were transfected into cells with or without the self-replicating RNA, and samples harvested after 16 h at 30°C were titered. Plaques were observed only in the samples harvested from cotransfected cells. Virus particles obtained from individual plaques were used to infect new monolayers, and the viral RNA patterns produced by the infected cells were analyzed. Crosses between TSCATts6 and several of the deletion mutants produced recombinants in which the 49S and 26S RNAs were indistinguishable from authentic 49S and 26S RNAs (Table 1).

The recombinants between TRCAT and CR3.4 gave a more complex result. Two examples of the viral RNAs synthesized in cells infected with stocks prepared from independent plaques are seen in Fig. 4A, lanes 1 and 2. In lane 1, there are four major viral RNA bands: two identical in size to viral 49S and 26S RNAs, one larger than 49S RNA, and one (probably not easily visualized in the photograph) only just slightly larger than 26S RNA. In lane 2, the most prominent RNA species are the one larger than 49S RNA, one the same size as 26S RNA, and one calculated to be 6.5 kb. An RNA of 6.5 kb is also present as a minor species in lane 1. The virus particles that gave rise to the pattern seen in lane 2 were subjected to a second round of plaque purification (based on the incorrect assumption that the 6.5-kb RNA might be a DI RNA). Several plaques were analyzed as described previously (Fig. 4B). RNA patterns similar to that produced by the original stock were observed (lanes 3 to 5) as well as ones that suggested that evolution of the RNA was occurring. The presence of the 6.5-kb RNA always correlated with the presence of an RNA larger than 49S RNA; as seen in lane 2, a decrease in the 49S-like RNA was always accompanied by a loss of the 6.5-kb RNA and the appearance of a new species migrating just above 26S RNA. Lane 1 shows a more heterogeneous mixture of RNAs. Samples of the RNAs (Fig. 4B, lanes 2 and 4) were analyzed further by PCR.

Figure 5 shows the structures of the parental RNAs and the locations of the primers that were used to generate the PCR products. The product obtained from wild-type RNA (Fig. 4A, lane 3) was the expected size of 569 bp, that generated from the RNA sample shown in Fig. 4A, lane 2, was 2.5 kbp, and the one from the RNA sample shown in Fig. 4B, lane 2, was about 770 bp. The latter fragment was cloned into M13mp19, and a single plaque was isolated and sequenced. A diagram of the region of the 770-bp cloned fragment that was sequenced (Fig. 5) shows that this RNA also contained two subgenomic promoter regions: one from TRCAT and one from CR3.4. On the basis of the RNA analysis, we believe that this RNA evolved from a larger recombinant RNA even though there was no evidence from the sequence for multiple steps as had been seen in the TRCAT/DI(26S.1) recombinant.

We carried out several other crosses (summarized in Table 1). In each case, plaques were isolated and used to infect cells for isolation and analysis of the Sindbis virus-specific RNAs. We found no evidence for two subgenomic RNAs in the few plaques that we analyzed from crosses with TSCATts6 as one of the parents. RNAs obtained from the cross between TSCATts6 and CR4.10 and from the cross between TSCATts6 and CR2.4 were examined further by PCR analysis across the junction region. The size of the PCR products was identical to that derived from wild-type RNA, indicating that there were no large insertions in this region of the genome, but more subtle changes would not have been detected. It is also possible that the crosses with TSCATts6 occur in other regions of the genome or that the inserted region was unstable and had been lost in the plaque isolates we analyzed.

Recombination between a Sindbis virus genomic RNA mutationally altered in the structural gene, E2 and DI(26S)

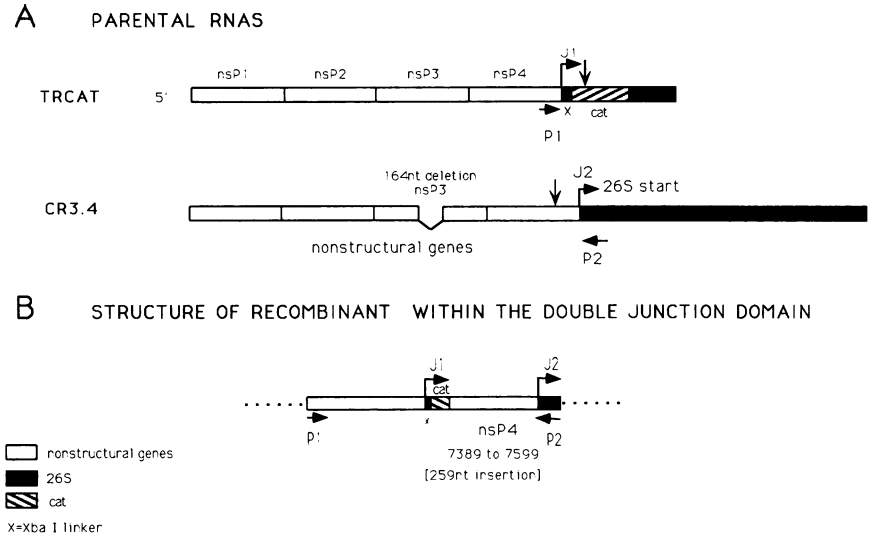

FIG. 5. Diagrams of parental RNAs TRCAT and CR3.4 (A) and of one recombinant in the region of the double-junction domain (B). Primers 1 and 2 (P1 and P2) were identical to those described in the legend to Fig. 1. Primer 2 was not present in TRCAT. The *Xba*I linker was derived from Toto1102 (28). The PCR product obtained by using these primers with wild-type Sindbis virus RNA should be 559 nucleotides (nt) in length. The recombinant yields a PCR product that is 259 nucleotides longer. The vertical arrows in the diagrams of TRCAT and CR3.4 indicate the points at which the two parental RNAs were joined in the recombinant RNA that was analyzed.

RNAs, or CR2.4. We also tested the ability of a 49S RNA with mutations in the E2 structural protein gene to undergo recombination with a DI RNA. The two parental RNAs are diagrammed in Fig. 6. Transfection of the two RNAs into cells gave rise to PFU. Stocks of virus were obtained from five plaques, and they all produced a pattern of viral RNAs identical to that shown in Fig. 7. We generated a PCR product by using the primers indicated in Fig. 6 and intracellular RNA isolated from cells infected with virus eluted from a single plaque. This PCR product was 1,413 nucleotides long, 384 nucleotides longer than the product synthesized with these primers and wild-type Sindbis virus RNA. The cDNA was cloned, and the sequence from one cDNA clone is diagrammed in Fig. 6. In addition to the now common occurrence of two junction regions, there was an insertion of 11 nucleotides almost immediately downstream of the initiating AUG of the capsid protein (position 7650 in Toto1000) and directly upstream of nucleotide 1118 of DI(26S) RNA. These nucleotides were not due to a PCR artifact but were present in the RNA; the product produced by primer extension on the RNA was the expected size, assuming that there was an 11-nucleotide insert. These 11 nucleotides, as a contiguous sequence, were not present in either the positive or negative strand of the Sindbis virus genome. Our search of the GenBank and EMBL data bases revealed a total of 15 genes or mRNAs that contain a perfect match for this stretch of 11 nucleotides. Three were bacterial genes, and two were herpes simplex virus genes. The remaining 10 were of diverse eukaryotic origin, and we had no basis for selecting any as a possible candidate to have donated the 11 nucleotides to the Sindbis virus genome. The AUG codon represented the start of an open reading frame that continued for a total of 31 amino acids. We have not determined whether a protein of the predicted size (3 kDa) was present in infected cells.

There was a second difference in this recombinant: the level of the 26S RNA produced in infected cells was significantly less than that of the larger subgenomic RNA. This

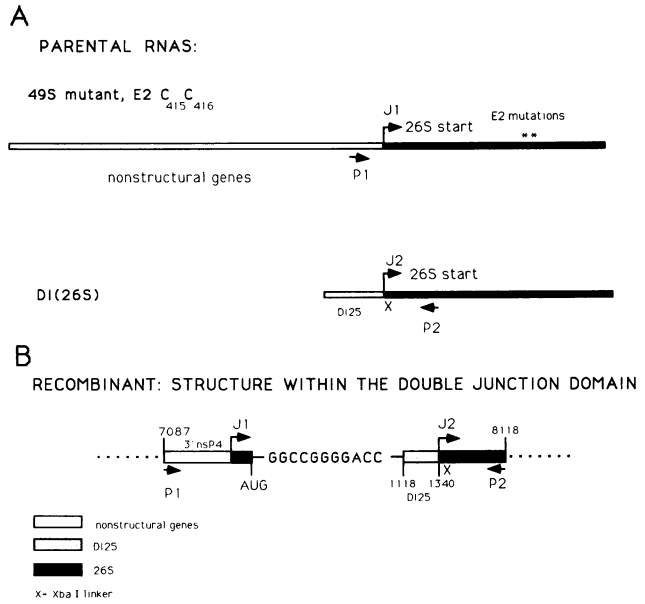

FIG. 6. Diagrams of parental RNAs $E2C_{415}C_{416}$ and DI(26S) (A) and of one recombinant in the region of the double-junction domain (B). The positive-sense primer (P1) lies at the 3′ end of nsP4 and is identical to that described in the legend to Fig. 1. The negative-sense primer (P2) lies within the capsid gene and is an 18-nucleotide oligomer derived from nucleotides 8099 to 8116 in Toto1100. DI(26S) contains an *Xba*I linker insertion in the noncoding 5′ 26S domain. $E2C_{415}C_{416}$ lacks this *Xba*I linker. The PCR product generated with these primers and wild-type Sindbis virus RNA is 1,029 nucleotides in length. The recombinant yields a product that is 384 nucleotides larger than wild type and contains an 11-nucleotide G+C-rich sequence of unknown origin that lies downstream of the first capsid AUG (at position 7650 in Toto1100) and upstream of nucleotide 1118 of DI(26S) RNA.

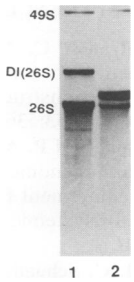

FIG. 7. RNA pattern observed for a single recombinant plaque isolate resulting from a cross involving the mutant $E2C_{415}C_{416}$ and DI(26S) RNAs. Recombinant virus released from cells transfected with the E2 mutant and DI(26S) was plaque purified. Virus from a single plaque was used to infect CEF monolayers in the presence of dactinomycin and [^{3}H]uridine. Lanes: 1, reference showing the positions of migration of the genomic parental RNAs and 26S RNA; 2, RNA pattern for cells infected with virus isolated from a single plaque.

contrasted with the other recombinants with two subgenomic RNA promoters. In those cases, 26S RNA was always the predominant of the two subgenomic RNA species. It did correspond, however, with the location of nucleotides derived from an *Xba*I linker. Previous studies had shown that the insertion of an *Xba*I linker at a position 14 nucleotides downstream from the start of the 26S cDNA led to a decrease in 26S RNA synthesis in infected cells (28). The diagram of the recombinant shown in Fig. 6 shows that the upstream promoter for 26S RNA synthesis came from the E2 mutant, which does not have the nucleotides derived from the *Xba*I linker. The downstream promoter derived from DI(26S) contains the six-nucleotide *Xba*I linker insertion.

We did not detect any recombination between CR2.4 and CR4.10, two RNAs with deletions in the nonstructural protein genes. However, we did observe recombination between CR2.4 and the E2 mutant (Table 1). These two RNAs are essentially complete full-length genomic RNAs, and neither contained foreign sequences. Cells infected with recombinant virus from this cross produced two subgenomic RNAs, indicating that both subgenomic RNA promoters were present in the recombinant RNA.

DISCUSSION

Sindbis virus can now be included among the positive-strand RNA viruses that undergo recombination. Mechanisms of recombination between RNA molecules have not been analyzed in detail, but template switching or copy choice is considered to be the most likely mechanism (15). Studies with poliovirus have provided evidence for a copy choice mechanism in which the viral polymerase switches templates during synthesis of the negative strand (17).

The five Sindbis virus recombinant RNAs that we analyzed by sequencing all had undergone recombinational events that introduced additional sequences into the 49S RNA genome. One of the most striking results was that many of the recombinant RNAs contained two functional subgenomic RNA promoters. This result, which at first seemed surprising, may be a consequence of the two-domain structure of the Sindbis virus genome. The 5′ domain encompasses the nonstructural protein genes, but the complete coding sequence of the nsP4 gene extends the domain one nucleotide plus the stop codon into the 26S RNA sequence (30). The 3′ domain codes for the structural proteins. It must also include the subgenomic RNA promoter for these genes to be expressed, making the subgenomic RNA promoter a component of both domains. In the wild-type genomic RNA, 48 nucleotides separate the two coding regions, but some of these noncoding nucleotides are important for the activity of the subgenomic RNA promoter. The minimal subgenomic RNA promoter is only 19 nucleotides upstream and 5 nucleotides downstream of the junction, but inclusion of the additional sequences surrounding the promoter enhances transcription of the subgenomic RNA (20, 27).

It has been possible to put large inserts in the region between the domains by constructing Sindbis virus cDNAs in which a foreign gene under the control of the subgenomic RNA promoter was positioned upstream of the 26S RNA and its promoter (10, 27). RNAs transcribed from these DNAs give rise to infectious virus particles that produce two subgenomic RNAs in infected cells. The recovery of such viruses provided the first demonstration that the genome could tolerate an expansion of the region between the domains and was instrumental in helping us understand the origin of the extra RNA species in our analysis of recombinant viruses.

Recombination between Sindbis virus RNAs occurring within a domain almost certainly would have to be precise to conserve coding or *cis*-acting regulatory sequences. In a recombination in which each parental RNA contributed one of the domains, recombinational events that created deletions, insertions, or rearrangements could occur anywhere outside of the functional domain and still give rise to an infectious recombinant. The length of the region between the two functional domains would depend on where recombination had occurred and on the stability of the recombinant RNA. The Sindbis virus genomes that were larger than the 49S RNA were unstable and eventually evolved to the same size as wild-type RNA.

The sequence of the recombinant RNAs and the analysis of the viral RNAs in infected cells indicated that those recombinants with two subgenomic RNA promoters did not arise by a simple crossover event between the original parental RNAs. In the cross between TRCAT and DI(26S.1), we could not distinguish whether the deletion in the CAT gene or the rearrangement of the E1 sequences (Fig. 1B) occurred before or after the recombinational event giving rise to the RNA that we analyzed. The recombinant between TRCAT and CR3.4, however, initially produced a second subgenomic RNA that was 6.5 kb, but the genomic RNA evolved, as shown by the change in its size as well as in the size of the subgenomic RNA (Fig. 4). In some of the crosses, the recombinant viruses produced 49S and 26S RNAs in infected cells that were identical in size to wild-type RNAs. These viruses may be recombinants that arose by recombination elsewhere in the genome, but they may also be ones that had evolved and lost a detectable insert.

In contrast to recombination between Sindbis virus RNAs, recombination between polioviruses (17) and between coronaviruses (3, 18) had to be precise for viable progeny to be detected. This difference may reflect differences in the organization of the genomes and the strategies of replication. The genome of poliovirus has one long open reading frame, and there would be little or no flexibility within this coding region for insertions or deletions to occur. In coronaviruses, leader-primed transcription may play a role in recombination (13, 22).

Both polioviruses and coronaviruses appear to undergo recombination at a much higher frequency than do Sindbis

viruses. The former two undergo easily detected recombination in cultured cells under conditions in which the parental viruses are also capable of replicating, a condition under which recombination was not observed between Sindbis viruses (5). There may be some mechanism(s) for template alignment that facilitates recombination between poliovirus (17) or between coronavirus RNAs but not between Sindbis virus RNAs. Recombination between Sindbis virus RNAs may be more analogous to that observed between the RNAs of bromoviruses, for which both homologous and nonhomologous recombination have been reported (2, 4). Sequence similarities in the nonstructural proteins of Sindbis virus and the bromoviruses had led to the proposal that these viruses evolved from a common ancestor (1).

King has pointed out that DI RNAs derived from RNA viruses are generated by nonhomologous recombinational events (15). Naturally occurring DI RNAs of Sindbis virus, and the closely related Semliki Forest virus, contain repeats and rearrangements of the genomic RNA (19, 25). It is likely that many of these DI RNAs arose by intermolecular recombination. One obvious example is a DI RNA, which we described several years ago, in which 142 nucleotides from the 5′ end of 26S RNA (with nucleotides 25 to 66 of the 26S RNA deleted) are joined directly to the 5′ terminus of the virion RNA (33).

The recombinant RNA derived from the $E2C_{415}C_{416}$ mutant and DI(26S) contained a stretch of 11 nucleotides that were not present in the Sindbis virus RNAs and may have a cellular origin. If so, it will represent the second time that we have observed the incorporation of cellular sequences into a Sindbis virus RNA. The first time was the addition of $tRNA^{Asp}$ to the 5′ terminus of DI RNAs (24). There are several other examples in which cellular sequences have been identified in viral RNAs: a ubiquitin-coding sequence was found in a bovine diarrhea virus (23), and a sequence from 28S rRNA was inserted into the hemagglutinin gene of an influenza virus (14). These findings provide support for the concept that the acquisition of cellular sequences is an important factor in the evolution of RNA viruses (32, 36).

This report is the first documentation of recombination between Sindbis virus RNAs in cultured cells, but there is also evidence that recombination within this genus of viruses had occurred in nature. Sindbis virus is one of the alphaviruses, a group of mosquito-borne viruses that belong to the *Togaviridae* family. Hahn et al. compared the sequence of Sindbis virus with that of two other alphaviruses, eastern equine encephalitis virus (EEEV) and western equine encephalitis virus (WEEV) (11). The sequences of the glycoproteins E2 and E1 of WEEV are closely related to those of Sindbis virus, whereas the rest of the genome is more similar to that of EEEV. Based on their analysis, Hahn et al. concluded that WEEV arose by recombination between a Sindbis-like virus and an EEE-like virus.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI11377 from the National Institute of Allergy and Infectious Diseases.

We thank Rebecca Wright for invaluable technical assistance, Ute Geigenmuller-Gnirke for providing the $E2C_{415}C_{416}$/DI(26S) recombinant and for her insight and comments on this work and the manuscript, and both Peter Bredenbeek and Charlie Rice for their comments on the manuscript. We thank K. G. Nitschko and M. Schlesinger for providing the Toto $E2C_{415}C_{416}$ plasmid and Charlie Rice, Mary Jones, Guangpu Li, and Julie Lemm for their generous gifts of plasmids and primers. We also thank the anonymous reviewers for their helpful comments on the manuscript.

REFERENCES

1. **Ahlquist, P., E. G. Strauss, C. M. Rice, J. H. Strauss, J. Haseloff, and D. Zimmern.** 1985. Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses. J. Virol. **53**:536–542.
2. **Allison, R., C. Thompson, and P. Ahlquist.** 1990. Regeneration of a functional RNA virus genome by recombination between deletion mutants and requirement for cowpea chlorotic mottle virus 3a and coat genes for systemic infection. Proc. Natl. Acad. Sci. USA **87**:1820–1824.
3. **Baric, R. S., K. Fu, M. C. Schaad, and S. A. Stohlman.** 1990. Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. Virology **177**:646–656.
4. **Bujarski, J. J., and P. Kaesberg.** 1986. Genetic recombination between RNA components of a multipartite plant virus. Nature (London) **321**:528–531.
5. **Burge, B. W., and E. R. Pfefferkorn.** 1966. Complementation between temperature-sensitive mutants of Sindbis virus. Virology **30**:214–223.
6. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85**:8998–9002.
7. **Geigenmuller-Gnirke, U., B. Weiss, R. Wright, and S. Schlesinger.** 1991. Complementation between Sindbis viral RNAs produces infectious particles with a bipartite genome. Proc. Natl. Acad. Sci. USA **88**:3253–3257.
8. **Grady, L. J., and W. P. Campbell.** 1989. Amplification of large RNAs (>1.5kb) by polymerase chain reaction. Biotechniques **7**:788–790.
9. **Grakoui, A., R. Levis, R. Raju, H. V. Huang, and C. M. Rice.** 1989. A *cis*-acting mutation in the Sindbis virus junction region which affects subgenomic RNA synthesis. J. Virol. **63**:5216–5227.
10. **Hahn, C., and C. Rice.** Unpublished results.
11. **Hahn, C. S., S. Lustig, E. G. Strauss and J. H. Strauss.** 1988. Western equine encephalitis virus is a recombinant virus. Proc. Natl. Acad. Sci. USA **85**:5997–6001.
12. **Hirst, G. K.** 1962. Genetic recombination with Newcastle disease virus, polioviruses, and influenza. Cold Spring Harbor Symp. Quant. Biol. **27**:303–308.
13. **Keck, J. G., S. A. Stohlman, L. H. Soe, S. Makino, and M. M. C. Lai.** 1987. Multiple recombination sites at the 5′ end of the murine coronavirus RNA. Virology **156**:331–341.
14. **Khatchikian, D., M. Orlich, and R. Rott.** 1989. Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. Nature (London) **340**:156–157.
15. **King, A. M. Q.** 1988. Genetic recombination in positive strand RNA viruses, p. 150–185. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics II. CRC Press, Inc., Boca Raton, Fla.
16. **King, A. M. Q., D. McCahon, W. R. Slade, and J. W. I. Newman.** 1982. Recombination in RNA. Cell **29**:921–928.
17. **Kirkegaard, K., and D. Baltimore.** 1986. The mechanism of RNA recombination in poliovirus. Cell **47**:433–443.
18. **Lai, M. M. C., R. S. Baric, S. Makino, J. Keck, J. Egbert, J. L. Leibowitz, and S. S. Stohlman.** 1985. Recombination between nonsegmented RNA genomes of murine coronaviruses. J. Virol. **56**:449–456.
19. **Lehtovaara, P., H. Soderlund, S. Keränen, R. F. Pettersson, and L. Kääriäinen.** 1982. Extreme ends of the genome are conserved and rearranged in the defective interfering RNAs of Semliki Forest virus. J. Mol. Biol. **156**:731–748.
20. **Levis, R., S. Schlesinger, and H. V. Huang.** 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. J. Virol. **64**:1726–1733.
21. **Levis, R., B. G. Weiss, M. Tsiang, H. Huang, and S. Schlesinger.** 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. Cell **44**:137–145.

22. **Makino, S., J. G. Keck, S. A. Stohlman, and M. M. C. Lai.** 1986. High frequency RNA recombination of murine coronaviruses. J. Virol. **57**:729–737.
23. **Meyers, G., N. Tautz, E. J. Dubovi, and H.-J. Thiel.** 1991. Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. Virology **180**:602–616.
24. **Monroe, S. S., and S. Schlesinger.** 1983. RNAs from two independently isolated defective interfering particles of Sindbis virus contain a cellular tRNA sequence at their 5′ ends. Proc. Natl. Acad. Sci. USA **80**:3279–3283.
25. **Monroe, S. S., and S. Schlesinger.** 1984. Common and distinct regions of defective-interfering RNAs of Sindbis virus. J. Virol. **49**:865–872.
26. **Ou, J.-H., C. M. Rice, L. Dalgarno, E. G. Strauss, and J. H. Strauss.** 1982. Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. Proc. Natl. Acad. Sci. USA **79**:5235–5239.
27. **Raju, R., and H. V. Huang.** 1991. Analysis of Sindbis virus promoter recognition in vivo using novel vectors with two subgenomic mRNA promoters. J. Virol. **65**:2501–2510.
27a. **Rice, C., et al.** Unpublished data.
28. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. J. Virol. **61**:3809–19.
29. **Sanger, F., S. Nicklen, and A. R. Coulsen.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
30. **Strauss, E. G., C. M. Rice, and J. M. Strauss.** 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. Virology **133**:92–110.
31. **Strauss, E. G., and J. H. Strauss.** 1986. Structure and replication of the alphavirus genome, p. 35–82. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Press, New York.
32. **Strauss, J. H., and E. G. Strauss.** 1988. Evolution of RNA Viruses. Annu. Rev. Microbiol. **42**:657–683.
33. **Tsiang, M., S. S. Monroe, and S. Schlesinger.** 1985. Studies of defective interfering RNAs of Sindbis virus with and without $tRNA^{Asp}$ sequences at their 5′ termini. J. Virol. **54**:38–44.
34. **Weiss, B., H. Nitschko, I. Ghattas, R. Wright, and S. Schlesinger.** 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. J. Virol. **63**:5310–5318.
35. **Xiong, C., R. Levis, P. Shen, S. Schlesinger, C. M. Rice, and H. V. Huang.** 1989. Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. Science **243**:1188–1191.
36. **Zimmern, D.** 1988. Evolution of RNA viruses, p. 211–240. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics II. CRC Press, Inc., Boca Raton, Fla.