JOURNAL OF VIROLOGY, Jan. 1982, p. 153–162
0022-538X/82/010153-10$02.00/0

Vol. 41, No. 1

Sequence Analysis of cDNA's Derived from the RNA of Sindbis Virions and of Defective Interfering Particles

STEPHAN S. MONROE,[1] JING-HSIUNG OU,[2] CHARLES M. RICE,[2] SONDRA SCHLESINGER,[1]* ELLEN G. STRAUSS,[2] AND JAMES H. STRAUSS[2]

Department of Microbiology and Immunology, Washington University School of Medicine, St. Louis, Missouri 63110,[1] and Division of Biology, California Institute of Technology, Pasadena, California 91125[2]

Received 15 June 1981/Accepted 28 August 1981

Sindbis virus generates defective interfering (DI) particles during serial high-multiplicity passage in cultured cells. These DI particles inhibit the replication of infectious virus and can be an important factor in the establishment and maintenance of persistent infection in BHK cells. In an effort to understand how these DI particles are generated and how they interfere with the replication of standard virus, we performed a partial sequence analysis of the RNA obtained from two independently isolated populations of DI particles and from two Sindbis virus variants and compared these with the RNA of the parental wild-type virus. The 3′-terminal regions of the RNAs were sequenced by the dideoxy chain terminating method. Internal regions of the RNA were examined by restriction endonuclease digestion of cDNA's made to the various RNAs and by direct chemical sequencing of 5′ end-labeled restriction fragments from cDNA made to the DI RNAs. One of the variant viruses examined was originally derived from cells persistently infected with Sindbis virus for 16 months and is resistant to interference by the DI strains used. In the 3′-terminal region of the RNA from this variant, only two base changes were found; one of these occurs in the 20-nucleotide 3′-terminal sequence which is highly conserved among alphaviruses. The DI RNA sequences were found to have been produced not by a single deletional event, but by multiple deletion steps combined with sequence rearrangements; all sequences examined are derived from the plus strand of Sindbis virion RNA. Both DI RNAs had at least 50 nucleotides of wild-type sequence conserved at the 3′ terminus; in addition, they both contained conserved and perhaps amplified sequences derived from the non-26S region of the genome which may be of importance in their replication and interference ability.

Defective interfering (DI) particles are nonviable deletion mutants produced by serial high-multiplicity passage of most viruses in cultured cells. These particles are characterized by their ability to interfere effectively with the replication of the homologous standard virus. The mechanism by which deletions are generated may vary for different viruses, but those sequences in the parental virus genome which are essential for replication and packaging should be conserved in the defective nucleic acids. To develop an understanding of how DI particles are generated and how they interfere with the growth of infectious virus, it is necessary to determine the sequence relationships between the standard and DI genomes. Our laboratories have had a long-standing interest in alphavirus replication. There is now considerable information available about the primary structure of the RNAs from two alphaviruses, Sindbis virus (12, 14; J.-h. Ou and E. Strauss, unpublished data) and Semliki Forest virus (3, 4), making it feasible to begin to probe the sequences of the nucleic acids from their DI particles.

The alphavirus genome is a single-strand 49S RNA which is infectious and can be translated to produce the nonstructural proteins involved in RNA replication (17, 20). During the infection cycle, intracellular full-length minus-strand RNA is made and serves as a template for the synthesis of progeny genome RNA and a 26S mRNA. Both of these RNA products are capped at the 5′ end and polyadenylated at the 3′ end. The 26S RNA is the 3′ third of the genome and codes for the virion structural proteins. Passaging either Sindbis or Semliki Forest virus at high multiplicity leads to the formation of DI particles in which the RNA appears to undergo successive deletions until a limit size of 18 to 20S RNA is reached (5, 19). On the basis of oligonucleotide fingerprinting and nucleic acid hybridizations, a model for the structure of these DI RNAs has been proposed in which 3′- and 5′-terminal sequences of virion RNA are conserved

and internal sequences of the genome are progressively deleted (6, 8). Recently, however, Pettersson reported that the 5'-terminal sequence of a limit size population of Semliki Forest virus DI RNAs is heterogeneous and different from the standard sequence (13). These results suggest that the structure of alphavirus DI RNAs may be more complex than had originally been thought.

Because of their ability to limit productive infection, DI particles can be involved in the establishment and maintenance of persistent viral infections in cultured cells (7). Weiss et al. established a line of BHK cells persistently infected with Sindbis virus by infecting the cells with a preparation of this virus containing a high ratio of DI particles to standard virus (21). One month after the persistent infection was established the infectious virus [Sin-1(1)] released from the infected cells was characterized as a small-plaque, temperature-sensitive mutant which was slightly less sensitive than the wild-type virus to interference by the original DI particles. Sixteen months after the persistently infected cells were in culture, the infectious virus [Sin-1(16)] cloned from these cells had become resistant to interference by the original DI particles (22). This phenotypic alteration must reflect a specific change or changes in the genome of the virion which prevent interference from occurring.

In this report we describe the results obtained by examining sequences of the RNA from two independently isolated DI particle populations of Sindbis virus and from the two variants cloned from persistently infected cells.

MATERIALS AND METHODS

Materials. Deoxy- and dideoxyribonucleoside triphosphates, primers, and T_4 polynucleotide kinase were purchased from P-L Biochemicals, Inc. The $p(dT)_7rG$ primer was purified as described previously (12). Bacterial alkaline phosphatase was purchased from Bethesda Research Laboratories, Inc. Oligo(dT)-cellulose was purchased from Collaborative Research, Inc. [^{3}H]uridine was purchased from New England Nuclear, [α-^{32}P]deoxynucleoside triphosphates, from Amersham Searle, and [γ-^{32}P]ATP from ICN. Restriction enzymes were purchased from New England Biolabs. Avian myeloblastosis virus reverse transcriptase was a generous gift from J. W. Beard.

Isolation and purification of Sindbis virus RNAs. The derivation of virus strains and DI particles is summarized in Table 1. Virus preparations to be used for RNA isolation were obtained by one-cycle infections of secondary chicken embryo fibroblasts grown in glass roller bottles. Virus particles released into the medium were concentrated by centrifugation and purified by sucrose gradient centrifugation. The RNA was isolated by extraction with phenol and chloroform and purified by oligo(dT)-cellulose chromatography (9). The RNA from Sin-1(1) was also purified by velocity gradient centrifugation. All procedures were carried out with aseptic techniques. Samples of the purified RNA were analyzed by agarose gel electrophoresis after denaturation with glyoxal (1). The RNAs isolated from the two DI preparations were predominantly mixtures of DI RNAs (Fig. 1) and are referred to as DI RNA populations to reflect their heterogeneity. The RNAs isolated from Sin-1(1) and Sin-1(16) comigrated with wild-type Sindbis 49S RNA (Fig. 1).

Sequencing by dideoxy chain termination. The 3'-terminal sequences of the viral RNAs were determined by the dideoxy chain termination technique (16) with the use of reverse transcriptase and the $p(dT)_7rG$ primer as described previously (12). The reactions contained 50 mM Tris-hydrochloride, pH 8.3, 50 mM KCl, 8 mM $MgCl_2$, 0.4 mM dithiothreitol, 0.4 pmol of RNA, and 0.1 μg of primer in 20 μl. cDNA synthesis was initiated by the addition of 4 U of avian myeloblastosis virus reverse transcriptase. The chain termination reactions contained 5 μM [α-^{32}P]dATP (400 Ci/mmol), 50 μM of the other three unlabeled deoxyribonucleoside triphosphates, and 0.5 to 50 μM dideoxyribonucleoside triphosphates. All reactions were incubated for 20 min on ice followed by 1 h at 37°C and were terminated by the addition of 10 μl of 10 M urea containing 0.1 mM dATP and dyes. Samples were heated for 2 min at 90°C, quickly cooled on ice, and loaded onto thin sequencing gels (15).

Synthesis of cDNA. We followed the methods described by Rice and Strauss (14a). Briefly, cDNA was synthesized in reactions containing: Tris-hydrochloride, pH 8.3, 50 mM; KCl, 50 mM; $MgCl_2$, 8 mM; dithiothreitol, 2 mM; sodium pyrophosphate, 4 mM; $oligo(dT)_{12-18}$, 10 to 20 μg/ml; digested calf thymus DNA, 300 μg/ml; template RNA, 30 to 60 μg/ml; and avian myeloblastosis virus reverse transcriptase, 500 U/ml. Analytical reactions contained ~1 μM α-^{32}P-labeled and 0.2 mM of each unlabeled deoxyribonucleoside triphosphate in 20 μl. Preparative reactions contained 1 mM unlabeled triphosphates and 50 μCi of [^{3}H]dTTP in 200 μl. Reactions were incubated at 42.5°C for 1 h, stopped by adding EDTA to 25 mM, phenol extracted, treated with 0.1 M NaOH at 60°C for 1 h to remove template RNA, and neutralized. Preparative cDNA's were purified further on a 5-ml Bio-Gel A5M column. The excluded fractions were pooled, lyophilized, resuspended in water, and ethanol precipitated two times.

Restriction enzyme digestion of cDNA. Single-stranded cDNA's were cleaved by digestion with restriction enzymes *Hae*III, *Hha*I, *Rsa*I, and *Taq*I (14a). To prepare fragments for sequencing, the digests were treated with alkaline phosphatase, phenol extracted, and labeled at the 5' ends with 1 mCi of [γ-^{32}P]ATP and polynucleotide kinase. The products were resolved on 5% or 6% sequencing gels and localized by autoradiography. Fragments to be sequenced were excised from the gel, eluted in 0.6 M sodium acetate, 0.1 M Tris-hydrochloride, pH 8, 2 mM EDTA, and 25 μg of tRNA per ml for 40 to 65 h at room temperature. The eluted DNA was ethanol precipitated twice to remove salt and was lyophilized before sequencing.

DNA sequencing. The procedures for base-specific chemical cleavages were based on those of Maxam and Gilbert (11) as modified by Smith and Calvo (18). The four base-specific reactions used were C, C + T, G + A and G. Samples were denatured for 2 min at

TABLE 1. Source of Sindbis RNAs used for sequencing studies

Virus	Description	Reference
DI-1	Obtained after 16 undiluted passages of Sindbis virus in BHK cells	9
DI-2	Obtained after 18 undiluted passages of Sindbis virus in chicken embryo fibroblasts	2
Sin-1(1)	Cloned from BHK cells 1 month after the establishment of persistent infection	21
Sin-1(16)	Cloned from the BHK cells described above but 16 months after persistent infection was established	22
Wild type	This is the parent of the HR strain and was obtained originally from B. Burge	

90°C in 80% formamide, quickly cooled, and analyzed on thin sequencing gels.

Thin sequencing gels. The gels were essentially the same as those described originally by Sanger and Coulson (15). The gels were either 30 by 40 by 0.04 cm or 30 by 80 by 0.04 cm. The voltage (1,800 to 2,400 V for 80-cm gels) was adjusted during the run to maintain the temperature of the gels at ~50°C. After electrophoresis, gels were transferred to Whatman 3 MM paper and exposed to presensitized film with intensifying screens at −70°C (10).

RESULTS

Sequencing of the 3'-terminal regions of Sindbis virus RNAs. On the basis of T_1-resistant oligonucleotide fingerprinting of Sindbis DI RNAs, Dohner et al. identified a characteristic 3' oligonucleotide conserved in all of the DI RNAs (2). They were not able to ascertain, however, to what extent these RNAs were identical to the standard virion RNA at the actual 3' terminus. Furthermore, all of the DI RNAs examined had a deletion near the 3' terminus, raising the possibility that the deletions always occurred at the same site. We sequenced the 3'-terminal regions of two populations of DI RNAs to determine which sequences are conserved and where the deletions begin. We also sequenced the 3' termini of the RNA from the two Sindbis virus variants, Sin-1(1) and Sin-1(16). The latter virus is no longer sensitive to interference by the DI particles used to initiate the persistent infection (22). We wished to know whether there was any correlation between resistance to interference by DI particles and alterations at the 3' terminus of the RNA.

Chain termination methods were used to determine the nucleotide sequence at the 3' terminus of these RNAs. Since the first nucleotide adjacent to polyadenylate [poly(A)] is C for many alphaviruses (12), we carried out avian myeloblastosis virus reverse transcriptase reactions using $p(dT)_7rG$ as the specific primer. The first few nucleotides adjacent to the priming site were determined with only one or two deoxyribonucleoside triphosphates in the reverse transcriptase reactions. The sequence of the first five nucleotides adjacent to the 3' poly(A) of all the viral RNAs analyzed is 5' A-U-U-U-C-poly(A) 3' and is identical to that found for Sindbis 49S RNA (12).

The sequence information obtained from reactions containing dideoxynucleoside triphosphates is shown in Fig. 2. The 20-nucleotide stretch at the 3' end of the RNA adjacent to the poly(A) tract, which has been shown to be highly conserved in the alphaviruses (12; J.-h. Ou, unpublished data), is identical in the wild-type parental strain, the two DI populations, and Sin-1(1). Sin-1(16) shows a single base change in this region, a T to C change at position 7 of the cDNA. Sin-1(1) is identical to the wild type for at least the first 150 nucleotides, whereas Sin-

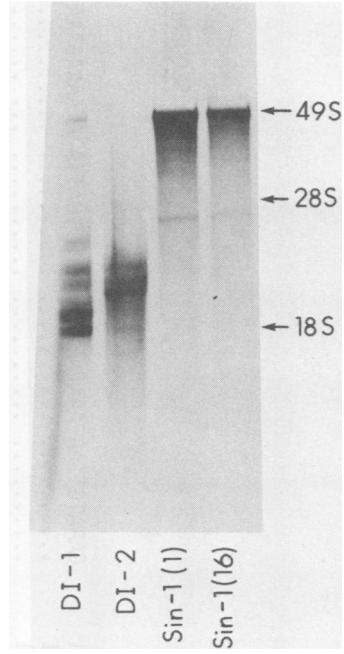

FIG. 1. Agarose gel electrophoresis of denatured Sindbis virus RNAs. Samples of [^{3}H]uridine-labeled RNAs were denatured in 10 mM sodium phosphate, pH 7, 1 M glyoxal, and 50% dimethyl sulfoxide for 1 h at 50°C. They were loaded on a horizontal 1% agarose gel (16 by 20 by 0.3 cm) made in 10 mM phosphate and electrophoresed for 3 h at 4 V/cm with constant buffer recirculation. The gel was impregnated with 2,5-diphenyloxazole (PPO), dried, and exposed to presensitized film at −70°C (10).

SEQUENCE

cDNA

26S (HR)	1	GAAATGTTAA AAACAAAATT TTGTTGATTA ATAAAAGAAA TAATAAAAGT TATGCAGACG CTGCGTGGCA TTATGCACCA	80
Wild Type	1	....	80
Sin-1(1)	1	....	80
Sin-1(16)	1	.....C... ....N.	80
DI-1	1	.... N..... N..... .N...N.... ...	73
DI-2	1	....N AGTCTGSATG AGCGGCCGGC TGACATNCCA	80

26S (HR)	81	CGCTTCCTCA GAAATACATT GAGTTTTTGG CGTCCGCTAG ATAAATGGTT AATATAGTGG TTATGTGGCA	150
Wild Type	81	....	150
Sin-1(1)	81	....	150
Sin-1(16)	81	....G.... .....N....	150
DI-2	81	AGGAGCCGNA GCATTTGPTT TTTGCTAGGA NCGGCAGAAG GTPATTTTGC ATGGTGTTGG NGTTCCT	147

FIG. 2. The 5′-terminal sequences of cDNA's transcribed from the 3′ end of Sindbis virus RNAs. The sequence of 26S cDNA was determined previously (14, 14a). In each case the 5′ terminal G is derived from the $p(dT)_7rG$ primer. The sequences were compared for homology by using a computer program. A dot indicates that the base is identical to the base listed for 26S cDNA. N, any nucleotide; P, purine; Y, pyrimidine; R, A or T; S, C or G.

1(16) shows, in addition to the change at position 7, a second change at position 136. DI-2 is identical to the wild type for the first 50 nucleotides, after which the sequences diverge. The following 54 nucleotides of DI-2 RNA (positions 51 to 104) are identical to positions 1410 to 1463 of Sindbis 26S RNA (HR strain) (14) and presumably to the parental wild-type 49S RNA. Positions 105 to 147 of DI-2 RNA are derived from the non-26S region of 49S RNA (E. Strauss, unpublished data). DI-1 RNA is identical to wild-type RNA for the first 73 nucleotides. After this, the DI-1 sequence becomes heterogeneous and could not be unambiguously determined. This indicates that at least some of the RNAs in the DI population diverge from the wild-type sequence at this point.

Restriction endonuclease digestion of viral cDNA's. An examination of the pattern of fragments obtained by digestion of DNA with restriction enzymes provides a means of comparing related nucleic acids. Several of the type II restriction enzymes cleave single-strand DNA, making it possible to apply this mapping procedure to cDNA copies of RNA. Accordingly, cDNA's were synthesized with reverse transcriptase by using the genomic RNAs as templates and a mixture of oligo(dT)$_{12-18}$ and digested calf thymus DNA as primers. After removal of the template, the cDNA's were digested with three different restriction enzymes, and the products were resolved on 5% sequencing gels (Fig. 3). Included in Fig. 3 is the pattern obtained from cDNA to 26S RNA to identify those fragments from this region of the genome.

The restriction patterns of the cDNA's from the two Sindbis variants [Sin-1(1) and Sin-1(16)] resemble quite closely those of Sindbis wild-type virion cDNA; the few differences are marked by arrows in Fig. 3. Sin-1(1) has one missing *Taq*I fragment derived from the non-26S region of the genome. Sin-1(16) shows alterations in two *Rsa*I fragments derived from 26S RNA, as well as one *Rsa*I fragment and four *Taq*I fragments derived from the non-26S RNA region of the RNA. We note, however, that many of the fragments observed result from partial (incomplete) cleavage (14a), and it is possible that a single change in the RNA could result in an observable change in more than one band.

The restriction patterns from the two DI cDNA's are quite different from those of the wild-type virus. Most of the wild-type restriction fragments and especially those from the 26S region are missing in the DI patterns. In addition, the two DI cDNA's have very different restriction patterns, although there are several bands in common (*Hae*III fragments C and D, for example). We note also that the patterns from the DI cDNA's are relatively simple, with few prominent bands. This limited complexity is similar to what was observed in the T_1-resistant oligonucleotide maps (2, 9).

Sequencing of *Hae*III fragments of DI-1 and DI-2. Preparations of cDNA to DI-1 and DI-2 were digested with *Hae*III, and the resulting fragments were labeled at the 5′ ends with T_4 polynucleotide kinase and [γ-^{32}P]ATP and were separated on a preparative gel. Fragments corresponding to those labeled in Fig. 3 were excised, eluted from the gel, and sequenced by the chemical methods of Maxam and Gilbert (11). The fragments varied in length from DI-1 fragment A, which is >850 nucleotides long, to fragment K, only 30 nucleotides long. (Fragment K is not shown in Fig. 3.) Some of the shorter fragments were sequenced in their entirety. For other fragments, either the first 35 nucleotides or the sequence beyond ~400 nucleotides from the 5′ end or both were not determined. The sequences obtained are shown in Fig. 4 and were aligned to maximize homology between fragments.

Figure 5A illustrates the location of 26S sequences in the DI cDNA's. This scheme incorporates data from both the dideoxy sequencing of the RNA from the 3′ terminus and from the chemical sequence of the internal restriction fragments. Three fragments from DI-1 contain 26S sequences. Fragment F is identical (except for a single base change at position 141) to the *Hae*III fragment of 26S RNA, 334 nucleotides long, located in the 26S sequence between nucleotide 1090 and nucleotide 1424 (14a). DI-1 fragment E and DI-1 fragment I, on the other hand, are composite fragments of 26S and non-26S sequence. Fragment E contains at its 5′ end part of the non-26S fragment 410 discussed in detail below (see also Fig. 5B), linked to 57 nucleotides found between nucleotides 1032 and 1089 of the 26S sequence, which is located directly adjacent to fragment F in 26S RNA. Fragment I, on the other hand, begins at its 5′ end with the 39 nucleotides of 26S sequence (nucleotides 1425 to 1463) adjacent to the other end of fragment F and terminates with 120 nucleotides of non-26S sequence which are common to fragment C1 from DI-1 (see Fig. 5B). There is no direct evidence that the three DI-1 fragments (E, F, and I) actually are joined together and are found in the same molecule, but the direct alignment with known 26S sequence suggests that they could be contiguous segments in the DI sequence.

Most of the remaining *Hae*III fragments which were sequenced contain some sequences in common, and the organization of these is shown in Fig. 5B. The top line shows two adjacent *Hae*III fragments from non-26S regions of the 49S virion RNA (called 410 and 450 here)

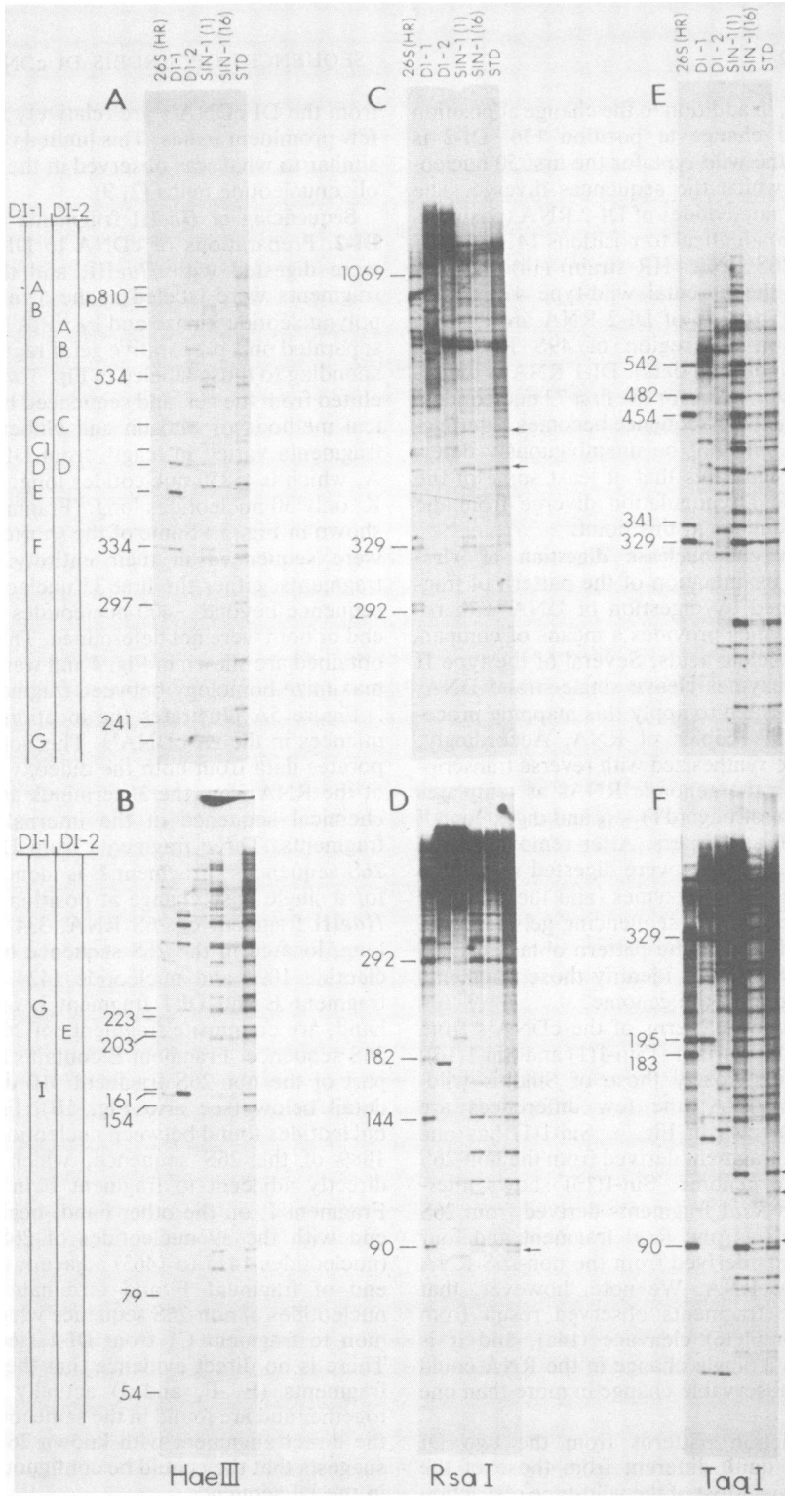

FIG. 3. Acrylamide gel analysis of restriction fragments of cDNA's transcribed from Sindbis virus RNAs. The conditions for synthesis of cDNA by use of random primers and subsequent cleavage with restriction enzymes are described in the text. Samples of the fragments were resolved on 40-cm 5% sequencing gels made in 100 mM Tris-borate, 2 mM EDTA, and 8.3 M urea. The gels were run at 40 W until the xylene cyanol dye had migrated 60 cm (upper panels) or 20 cm (lower panels). The lengths of fragments from 26S cDNA which were sequenced previously are indicated. *Hae*III fragments from the DI cDNA's which were subjected to direct sequencing are identified with letters. Fragment K ran off this gel and was identified from other gels. Arrows in each panel indicate fragments from wild-type Sindbis cDNA missing from variant cDNA.

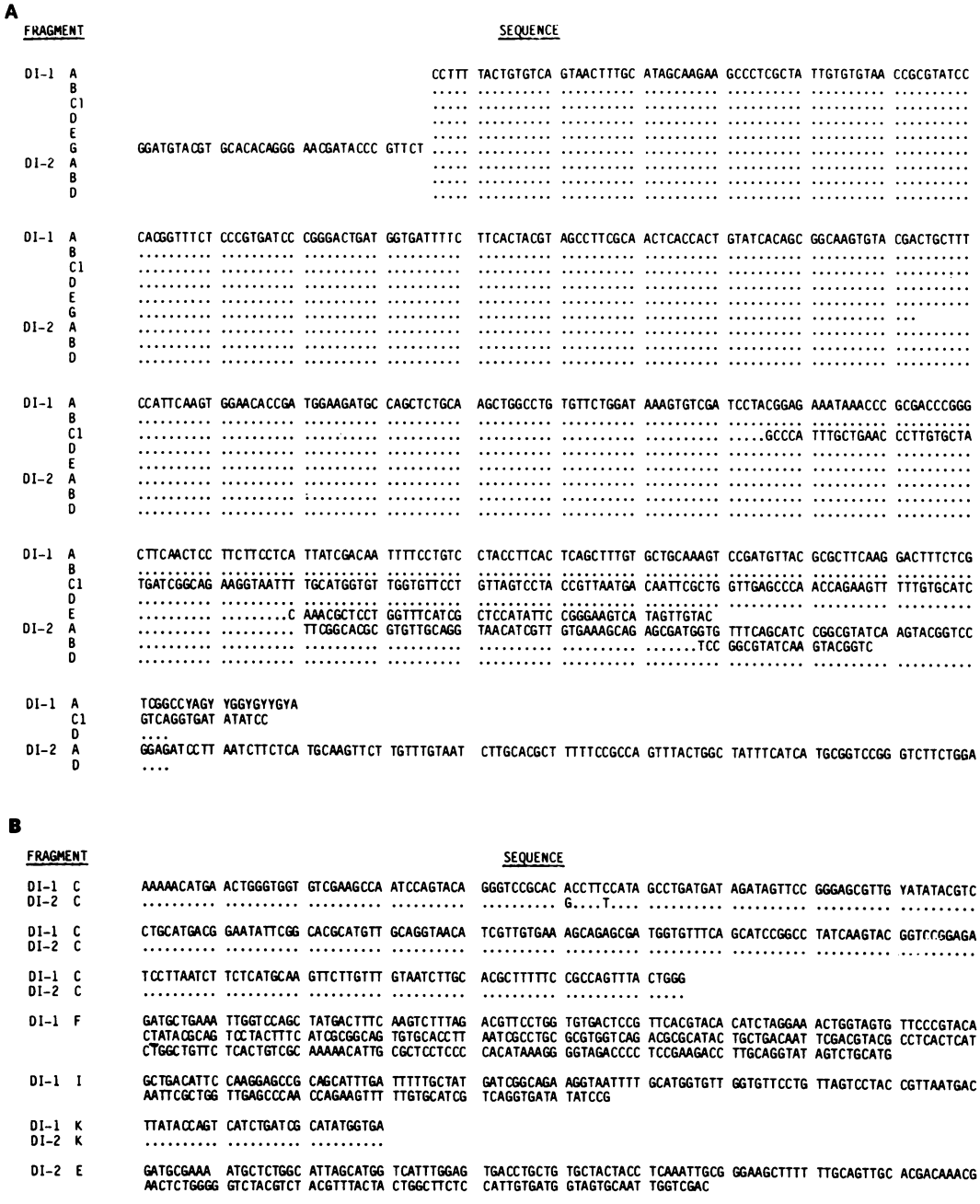

FIG. 4. 5′-Terminal sequences of restriction fragments generated by *Hae*III cleavage of DI cDNA's. The fragments are lettered as indicated in Fig. 3. The sequences were determined by chemical cleavage and analysis on acrylamide gels (11, 18). (A) The sequences of fragments with shared terminal regions are aligned to show the extent of homology. A dot indicates that the base is identical to the base listed for DI-1 fragment A. The ambiguous letter code is described in the legend for Fig. 2. Sequences of DI-1 fragments A, B, C1, D, and E and DI-2 fragments A, B, and D were determined beginning approximately 35 bases from the 5′ end. (B) The sequences of the remaining fragments analyzed from the DI cDNA's. The single base change between 26S cDNA and DI-1 fragment F is underlined.

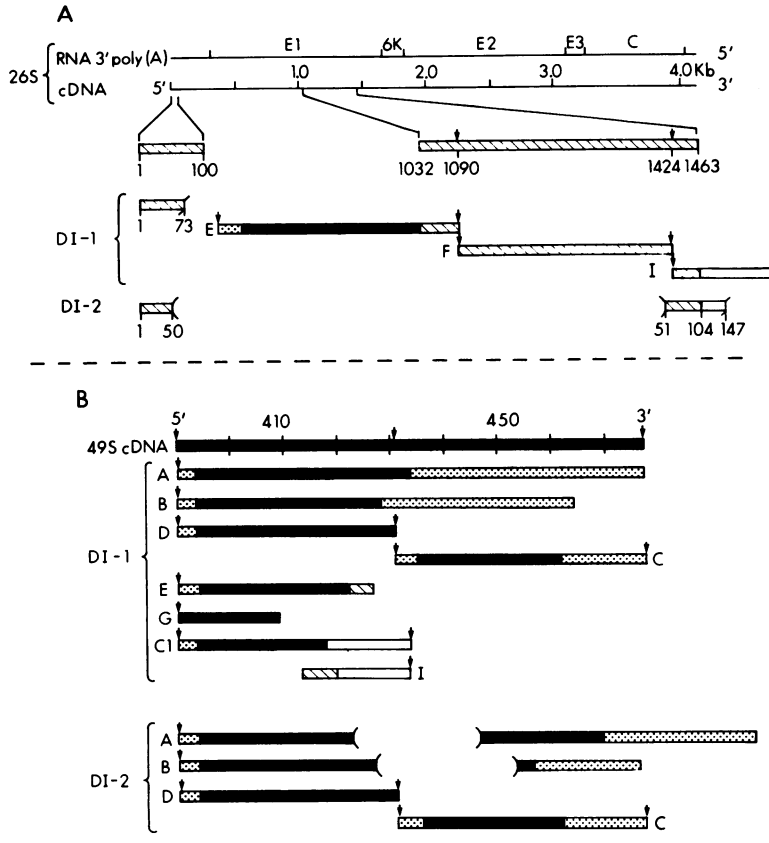

FIG. 5. Model for the sequence organization of Sindbis DI cDNA's. (A) Organization of the 5' region of the DI cDNA's based on chain termination sequencing and analysis of *Hae*III fragments. The 26S RNA and cDNA are shown at the top of the figure. Sequences are numbered from the 5'-terminal G residue of the cDNA. *Hae*III fragments from DI-1 cDNA are identified by letters. The order of the structural protein genes in 26S RNA is as follows: C, capsid; E3, E2, and E1, viral glycoproteins (17). (B) Organization of several internal *Hae*III fragments of DI cDNA which contain sequences in common. At the top are two contiguous *Hae*III fragments, 410 and 450 nucleotides in length, respectively, present in the non-26S region of HR 49S cDNA. Several *Hae*III fragments of cDNA to DI-1 and DI-2 RNAs are diagrammed below. The following designations are used for both A and B: cross-hatched bars, sequences found in 26S cDNA; solid bars, sequences in the nonstructural (non-26S) part of 49S cDNA homologous with fragments 410 and 450; open bars, other non-26S sequences; stippled bars, regions of DI fragments for which sequence data were not obtained. The vertical arrows indicate *Hae*III cleavage sites in the cDNA.

whose sequence is known (E. Strauss, unpublished data) but whose location in the genome is unknown. The region defined by these two fragments is 860 nucleotides long or approximately 10% of the non-26S portion of the Sindbis genome. Fragments C and D (of both DI-1 and DI-2) are the same as 450 and 410, respectively, beginning and ending at legitimate *Hae*III sites, and they appear to represent unaltered sequences in this region. Fragment A of DI-1 is a partial digestion product containing both fragments (probably in their entirety). Four other of the most prominent *Hae*III fragments of DI-1 (fragments B, C1, E, and G) as well as two fragments from DI-2 (fragments A and B) begin with the same sequence (the 5' end of fragment 410). Fragment B of DI-1 is shorter than A and probably contains a deletion whose end points are unknown. Fragment C1 of DI-1 contains the common 5' sequence plus the same sequence found attached to 26S sequence in fragment I. This sequence is present at an unknown location in 49S cDNA (E. Strauss, unpublished data). Fragment E of DI-1 is a rearrangement of non-26S and 26S sequences as described above. DI-2 fragments A and B are simple deletions of 241 and 245 nucleotides, respectively.

The other fragments which were sequenced, E from DI-2 and K from both DI populations, contain sequences found in the non-26S regions

of the 49S sequence (E. Strauss, unpublished data). Thus, all of the DI sequence obtained so far can be identified in the 49S RNA of Sindbis virus, and no host or nonviral sequences have been found. This is consistent with earlier studies with nucleic acid hybridization which showed that Sindbis DI RNA can be completely protected from RNase by hybridization to RNA complementary to virion 49S RNA (6).

DISCUSSION

The 3'-terminal sequences of the RNAs of several different alphaviruses show strong conservation of the first 20 nucleotides, suggesting that there are evolutionary pressures resulting in conservation of this region (12; J.-h. Ou, unpublished data). This 3'-terminal conserved sequence was also found in two DI RNA populations of Sindbis virus and two variants of this virus obtained from persistently infected BHK cells, supporting the conclusion that these sequences play an essential role in virus replication.

The variant [Sin-1(16)] from persistently infected cells which is resistant to interference by DI particles (derived from wild-type virus) had two base changes in the 3'-terminal sequence. One is at position 7 within the sequence which is most highly conserved among related alphaviruses, and the other is at position 136. If the first 20 nucleotides form a replicase binding site and if interference involves competition for replicase molecules, then the change from A to G at position 7 may be important for the phenotype of this variant. However, resistance to interference must be more complex, since variant virus cloned from cells after 1 month of persistent infection had 3' sequences identical to those of the wild-type parental strain, yet showed some resistance to interference.

The RNA obtained from the two DI strains had sequences identical to Sindbis HR or wild-type virus for at least the first 50 nucleotides adjacent to the 3' poly(A). Both DI RNA populations were heterogeneous, but since only one terminal sequence was obtained all of these species must contain the standard 3' sequence (within the limits of detection). The two DI RNA populations diverge from the wild-type sequences at different locations (Fig. 2), demonstrating that deletions can occur at more than one site during the generation and subsequent evolution of DI RNAs. Since DI-1 was generated by passing Sindbis virus in BHK cells and DI-2 was generated by passing the virus in chicken cells, it is possible that the host cell might influence the exact site or extent of deletions.

In the case of DI-2 the first (3'-terminal) deletion comprises 1,359 nucleotides and extends from position 51 through 1409 of the HR 26S RNA sequence. Following a stretch of 54 nucleotides (positions 51 to 104 of the DI RNA), there is a second deletion which jumps into non-26S sequence (Fig. 5A). In the case of DI-1, the 3'-terminal sequence was identical to the HR sequence until nucleotide 73 from the 3' poly(A), where it became heterogeneous. It will be necessary to sequence the individual DI RNA species to determine whether the heterogeneity is due to only part of the DI population diverging at position 73 and others diverging later or whether all diverge at position 73 but restart at different places in the wild-type genome. In this regard we note that both DI RNA populations lack a specific 3' oligonucleotide which we believe extends from nucleotides 117 to 148 of the wild-type sequence (2, 14). Thus, the entire DI-1 population must diverge by nucleotide 148, and in all cases the 3'-terminal deletion starts within the relatively narrow region between nucleotides 50 and 148 of the 26S RNA sequence.

It was surprising that so many of the *Hae*III fragments from both DI cDNA's contain different lengths of identical sequences. These sequences correspond to sequences identified in the nonstructural region of 49S RNA. They are not at the 5' end of the genome, but their exact location has not yet been determined (J.-h. Ou and E. Strauss, unpublished data). Some of these fragments must have been transcribed from different RNA molecules or from different regions of the same molecule. For example, fragments E and C1 from DI-1 cannot be derived from any of the larger fragments and fragments A, B, and D from DI-2 are distinct (Fig. 5B). The DI RNAs are heterogeneous, and it is possible that the fragments are derived from different DI RNA molecules. An alternative explanation—that they represent repeated sequences in the same molecule—is supported by the observation that repeated sequences are present in at least one Semliki Forest DI RNA (10a).

A comparison of sequences present in the two different DI cDNA's indicates that there may be some other regions of special significance. The DI-2 sequences from 66 to 147 and DI-1 fragment I are identical. They both leave the 26S sequence at the same nucleotide and enter the same non-26S sequence at the same nucleotide. This suggests that the two DI populations may follow a similar pathway of evolution. The same non-26S sequence (beginning, however, at a different point) is also found in DI-1 fragment C1 attached to the parental sequence of fragment 410 (Fig. 5B). Furthermore, two fragments diverge near nucleotide 321 of the wild-type sequence of fragment 410 (Fig. 5B). DI-1 fragment E diverges at this nucleotide to 26S sequence, indicating a rearrangement, whereas DI-2 fragment A diverges closely nearby in a simple

deletion. It is possible that some regions of the RNA may have a secondary structure which facilitates deletions.

In conclusion, the structure of the DI RNA and the relation of the DI sequence to the virion genome are more complex than was first envisaged. It is clear the DI RNA cannot arise solely from large deletions in the middle of the genome, retaining only the 3′ and 5′ ends. This result is in agreement with T_1 oligonucleotide studies which showed that Sindbis DI RNAs contain multiple deletions (2). However, the 3′ termini of both DI RNAs are conserved, and the entire DI RNA sequence appears to be derived from the plus strand of the viral 49S RNA. Whether the conserved termini represent replicase binding sites or whether they are important for a three-dimensional configuration essential for replication or encapsidation, or both, cannot be determined at this time. Similarly, there is conservation (and perhaps amplification) of one region of non-26S sequence and conservation of a short stretch of 26S sequence. It is possible that these conserved sequences could also be important for DI function. One possible hypothesis is that they could serve as binding sites for capsid protein during encapsidation. Further sequence studies on these and on related DI RNAs will be necessary to resolve these questions.

ACKNOWLEDGMENTS

We thank T. Hunkapiller for the computer program which made it possible to determine the sequence homologies reported here and Barbara Weiss for encouragement, critical comments, and help in preparing the RNA.

This research was supported by Public Health Service grants AI 11377 (to S.S.) and AI 10793 (to J.H.S.) from the National Institute of Allergy and Infectious Diseases, GM 06965 (to J.H.S.) from the National Institute of General Medical Science, and P30CA16217 (to S.S.) from the National Cancer Institute, and by grant PCM-8022830 (to J.H.S.) from the National Science Foundation.

LITERATURE CITED

1. **Carmichael, G. G., and G. K. McMaster.** 1980. The analysis of nucleic acids in gels using glyoxal and acridine orange. Methods Enzymol. **65**:380–391.
2. **Dohner, D., S. Monroe, B. Weiss, and S. Schlesinger.** 1979. Oligonucleotide mapping studies of standard and defective Sindbis virus RNA. J. Virol. **29**:794–798.
3. **Garoff, H., A.-M. Frischauf, K. Simons, H. Lehrach, and H. Delius.** 1980. Nucleotide sequence of cDNA coding for Semliki Forest virus membrane glycoproteins. Nature (London) **288**:236–241.
4. **Garoff, H., A.-M. Frischauf, K. Simons, H. Lehrach, and H. Delius.** 1980. The capsid protein of Semliki Forest virus has clusters of basic amino acids and prolines in its amino terminal region. Proc. Natl. Acad. Sci. U.S.A. **77**:6376–6380.
5. **Guild, G. M., and V. Stollar.** 1975. Defective interfering particles of Sindbis virus. 3. Intracellular viral RNA species in chick embryo cell cultures. Virology **67**:24–41.
6. **Guild, G. M., and V. Stollar.** 1977. Defective interfering particles of Sindbis virus. 5. Sequence relationships between SV_{STD} 42S RNA and intracellular defective viral RNAs. Virology **77**:175–188.
7. **Holland, J. J., S. I. T. Kennedy, B. L. Semler, C. L. Jones, L. Roux, and E. A. Grabau.** 1980. Defective interfering RNA viruses and the host-cell response, p. 137–192. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 16. Plenum Press, New York.
8. **Kennedy, S. I. T.** 1976. Sequence relationships between the genome and the intracellular RNA species of standard and defective-interfering Semliki Forest virus. J. Mol. Biol. **108**:491–511.
9. **Kennedy, S. I. T., C. J. Bruton, B. Weiss, and S. Schlesinger.** 1976. Defective interfering passages of Sindbis virus: nature of the defective virion RNA. J. Virol. **19**:1034–1043.
10. **Laskey, R. A.** 1980. The use of intensifying screens or organic scintillators for visualizing radioactive molecules resolved by gel electrophoresis. Methods Enzymol. **65**:363–371.

10a. **Lehtovaara, P., H. Söderlund, S. Keränen, R. F. Pettersson, and L. Kääriäinen.** 1981. 18S defective interfering RNA of Semliki Forest virus contains a triplicated linear repeat. Proc. Natl. Acad. Sci. U.S.A. **78**:5353–5357.

11. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
12. **Ou, J.-h., E. G. Strauss, and J. H. Strauss.** 1981. Comparative studies of the 3′-terminal sequences of several alphavirus RNAs. Virology **109**:281–289.
13. **Pettersson, R. F.** 1981. 5′-Terminal nucleotide sequence of Semliki Forest virus 18S defective interfering RNA is heterogenous and different from the genomic 42S RNA. Proc. Natl. Acad. Sci. U.S.A. **78**:115–119.
14. **Rice, C. M., and J. H. Strauss.** 1981. Nucleotide sequence of the 26S mRNA of Sindbis virus; deduced sequence of the encoded viral structural proteins. Proc. Natl. Acad. Sci. U.S.A. **78**:2062–2066.

14a. **Rice, C. M., and J. H. Strauss.** 1981. Synthesis, cleavage and sequence analysis of DNA complementary to the 26S messenger RNA of Sindbis Virus. J. Mol. Biol. **150**:315–340.

15. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87**:107–110.
16. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
17. **Schlesinger, M. J., and L. Kääriäinen.** 1980. Translation and processing of alphavirus proteins, p. 371–392. *In* R. W. Schlesinger (ed.), The togaviruses. Academic Press, Inc., New York.
18. **Smith, D. R., and J. M. Calvo.** 1980. Nucleotide sequence of the *E. coli* gene coding for dihydrofolate reductase. Nucleic Acids Res. **8**:2255–2274.
19. **Stark, C., and S. I. T. Kennedy.** 1978. The generation and propagation of defective-interfering particles of Semliki Forest virus in different cell types. Virology **89**:285–299.
20. **Strauss, J. H., and E. G. Strauss.** 1977. Togaviruses, p. 111–166. *In* D. P. Nayak (ed.), The molecular biology of animal viruses, vol. 1, Marcel Dekker, Inc., New York.
21. **Weiss, B., R. Rosenthal, and S. Schlesinger.** 1980. Establishment and maintenance of persistent infection by Sindbis virus in BHK cells. J. Virol. **33**:463–474.
22. **Weiss, B., and S. Schlesinger.** 1981. Defective interfering particles of Sindbis virus do not interfere with the homologous virus obtained from persistently infected BHK cells but do interfere with Semliki Forest virus. J. Virol. **37**:840–844.