Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 5235–5239, September 1982
Biochemistry

Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA

(Sindbis virus/RNA sequence determination/sequence homology/nonstructural proteins/transcription initiation)

JING-HSIUNG OU, CHARLES M. RICE, LYNN DALGARNO*, ELLEN G. STRAUSS, AND JAMES H. STRAUSS†

Division of Biology, California Institute of Technology, Pasadena, California 91125

Communicated by Ray D. Owen, June 7, 1982

ABSTRACT **The alphaviruses produce two mRNAs after infection: the genomic (49S) RNA which is translated into the nonstructural (replicase) proteins and the subgenomic (26S) RNA which serves as the mRNA for the virion structural proteins. The sequence of the region of the genomic RNA that contains the 5′ end of the subgenomic RNA and the 5′ flanking sequences in the genomic RNA were determined for several alphaviruses. A highly conserved sequence of 21 nucleotides was found which includes the first two nucleotides of the subgenomic RNA and the 19 nucleotides preceding it. We propose that the complement of this sequence in the minus strand is the recognition site used by the viral transcriptase for initiation of transcription of 26S RNA and that, in general, such short recognition sequences are commonly used among the RNA viruses. The COOH-terminal sequence of the nonstructural polyprotein precursor has been deduced for each virus. These protein sequences are highly homologous and are followed by multiple in-phase termination codons clustered in the nontranslated region of the 26S RNA in each case. In contrast to the proposed transcriptase recognition site, the particular triplets used for a given conserved amino acid have diverged markedly during evolution of these viruses. The protein homology is sufficient, however, for deduction of the correct coding phase of the RNA and allows the alignment of the corresponding nucleic acid sequence data from different alphaviruses without knowledge of the sequence of the entire genomes.**

The *Alphavirus* genus of the Togaviridae family includes more than 20 distinct viruses, many of which are important human or veterinary pathogens. These viruses are transmitted in nature via arthropod vectors and have the ability to replicate in a wide range of vertebrates including both avian and mammalian hosts. Because these viruses differ in geographical distribution, host range, and the pathological result of infection but are extremely similar in molecular architecture and in the pattern of events comprising viral replication, they provide an ideal system for studying the evolutionary relationships among members of a structurally and functionally related virus group. Viruses are known to mutate rapidly in general (1, 2), but the wide host range of this group of viruses and the alternation in nature between vertebrate and invertebrate hosts means that the alphaviruses are subject to different selective pressures compared to viruses with more limited host ranges.

The alphavirus genome consists of a single plus-strand RNA molecule about 13,000 nucleotides long with a sedimentation coefficient of 49 S (reviewed in ref. 3). This genomic RNA serves as the mRNA for the nonstructural proteins which function as the replicase/transcriptase activities required for replication of the viral RNA and for transcription of a 26S subgenomic mRNA encoding the virion structural proteins. The 26S RNA contains sequences identical to the 3′-terminal one-third of the genomic 49S RNA. The nucleotide sequences of the genomic RNA preceding and including the beginning of the 26S RNA (called the "junction region") would be expected to contain signals important in the termination of translation of the nonstructural polyprotein precursor as well as sequences important for 26S RNA transcription.

By comparing several different alphavirus sequences in a given region of interest, we have found that it is possible to define potentially important features involved in alphavirus replication by virtue of their conservation (or lack thereof) during evolution (4–7). Here we report the sequences of four different alphavirus junction regions. Comparison of these sequences has revealed several common features and provides insight into the translation and transcription of the 26S and 49S RNAs and the evolution of these viruses.

MATERIALS AND METHODS

Preparation of Vaccinia Guanylyltransferase. Guanylyltransferase was prepared from vaccinia virus (WR strain, isolate 11; a gift from W. K. Joklik) by a simplified version of the method of Paoletti *et al.* (8); 2 mg of vaccinia virus in 640 μl of 1 mM Tris·HCl (pH 9.0) was used for each preparation of the enzyme. After disruption of the virus cores, the preparation was briefly sonicated to shear viral DNA and then centrifuged at 136,000 × *g* for 60 min. The guanylyltransferase activity present in the supernatant was stable for 2 weeks at 4°C.

5′-End Labeling of Alphavirus 26S RNA. The 26S RNA of Sindbis virus (HR strain), Semliki Forest virus, and Middelburg virus were prepared as described (4) and decapped by β-elimination by the method of Rose and Lodish (9). Decapped RNA was then end-labeled by using conditions modified from Ahlquist *et al.* (10). The 100-μl reaction mixture contained 0.16 mM ATP, 1 mM $MgCl_2$, 1 mM dithiothreitol, 2.5 μM [α-^{32}P]GTP (410 Ci/mmol; 1 Ci = 3.7 × 10^{10} becquerels; Amersham), 50 mM Tris·HCl (pH 7.8), and 15 μl of guanylyltransferase extract. After incubation at 37°C for 15 min, the reaction was terminated by phenol/chloroform extraction followed by ethanol precipitation as described (4). RNA pellets were resuspended in 50 μl of 10 mM Tris·HCl, pH 8.2/10 mM NaCl/1 mM EDTA/0.1% $NaDodSO_4$ and purified from the unincorporated label by gel filtration on a Bio-Gel A-5m column. The excluded volume fractions containing the RNA were pooled and the RNA was ethanol precipitated twice with carrier RNA and resuspended in water to a final concentration of about 2,000 cpm/μl. More than 1 × 10^5 cpm was incorporated per 30 μg of RNA, which was sufficient for several experiments.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

* Present address: Biochemistry Dept., The Australian National University, P.O. Box 4, Canberra, Australia 2600.
† To whom reprint requests should be addressed.

Direct Enzymatic Sequence Determination of the RNA from the 5′ End. Partial RNase digestion of the 5′-end labeled RNA was performed by the method recommended by P-L Biochemicals (method E998) and the products were separated on sequencing gels. A brief description of the reaction conditions is in the legend of Fig. 1.

Single-Stranded cDNA Synthesis and Chemical Sequence Determination. The 49S RNAs of Sindbis, Middelburg, and Semliki Forest viruses were prepared as described (4). Ross River virus (T48 strain from R. Shope) 26S and 49S RNAs were prepared from infected BHK cells. Single-stranded cDNA to these RNAs were synthesized with reverse transcriptase and calf thymus DNA or oligo(dT) primers as described (11). Preparation, isolation, and chemical sequence determination of 5′-end-labeled restriction fragments from these cDNAs have been described (11).

RESULTS

In recent studies, nearly the entire 26S RNA nucleotide sequences for Sindbis virus (5) and Semliki Forest virus (12, 13) have been determined. However, the 5′ termini of these RNAs, a necessary landmark for defining the junction region, could not be localized unambiguously from these data. Thus, we began by directly determining the 5′-terminal 26S RNA sequences for several alphaviruses. The 26S RNAs of Sindbis, Middelburg, and Semliki Forest viruses were decapped by β-elimination, end-labeled with guanylyltransferase, and then partially digested by using endonucleases with four different base specificities. The resulting products were separated on acrylamide gels and the sequences were determined from the ladder produced. Fig. 1 presents part of the data to illustrate the technique. Each RNA was analyzed at least twice, and the sequences obtained are shown in lowercase letters in Fig. 2. In general this enzymatic method gives a clean signal for the purines, but the pyrimidines are often ambiguous (Fig. 1).

In order to determine the sequences preceding the start of 26S RNA and to clarify the direct RNA sequence analysis data, chemical sequence analysis of single-strand cDNA made to 49S RNA of Sindbis, Middelburg, Semliki Forest, and Ross River viruses was used (Fig. 2, uppercase letters). The method involved analysis of *Hae* III fragments of the cDNA (11) by the methods of Maxam and Gilbert (14). For Sindbis and Middelburg viruses, restriction fragments produced by *Hae* III digestion of cDNA to 49S RNA were randomly selected for analysis. Identification and alignment of the sequences in the junction region was done by computer. The RNA sequences and the possible protein sequences obtained by translation of the RNA in all three reading frames were searched by using homology routines, and the fragments that contained the start of 26S RNA and the start of the capsid protein were identified. These fragments end at nucleotide −126 for Sindbis virus and −180 for Middelburg virus (numbering begins at the start of 26S RNA) (Fig. 2). Furthermore, because the *Hae* III sites were found at different positions in the two viruses, adjacent *Hae* III fragments could be positioned by homology. This is possible because of the pronounced homology at the protein level among the various alphaviruses, and the use of homology routines often allowed corresponding fragments in any two viruses to be identified and aligned. Thus, the sequences of the underlined nucleotides in Fig. 2 (Middelburg virus, nucleotides −180 to −183; Sindbis virus nucleotides −126 to −129) were not determined directly but were inferred from the activity of *Hae* III and the homology between the two RNAs and their translated proteins.

For Semliki Forest virus, a simplified procedure was used to identify the junction fragment. cDNAs to both 26S and 49S RNA were made with an oligo(dT) primer and then digested with *Hae* III, and the resulting fragments were compared on gels. Because of premature termination by the reverse transcriptase, the complexity of the 49S pattern was only slightly greater than that of the 26S RNA pattern. From our direct 5′-end sequence analysis and the previously published 26S RNA sequence (12) it was known that the first *Hae* III site in 26S RNA was 83 nucleotides from the cap. Only 12 49S RNA specific fragments larger than 83 nucleotides were found, and these were excised and scanned by chemical sequence analysis with the

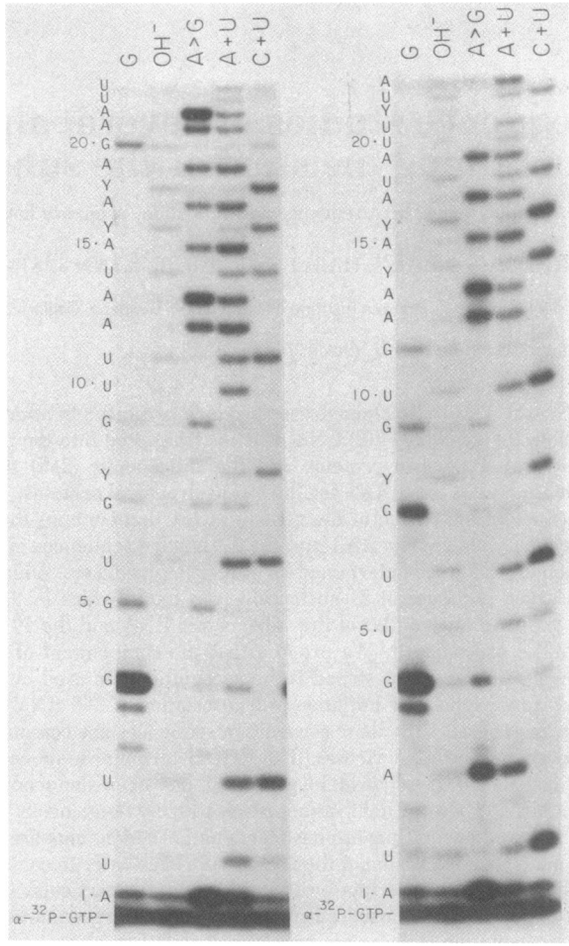

FIG. 1. Direct 5′-end sequence analysis of the 26S RNAs of Semliki Forest virus (*Left*) and Middelburg virus (*Right*). About 2,000 cpm of end-labeled RNA was enzymatically digested at 56°C for 3–10 min (depending on the degree of digestion preferred). Alkaline hydrolysis was at 90°C for 6 min. G reaction: 1 μl of RNA was added to 3 μl of buffer I containing 1 unit of RNase T1. A>G reaction: 1 μl of RNA was added to 3 μl of buffer II containing 1 unit of RNase U2. A+U reaction: 1 μl of RNA was added to 3 μl of buffer I containing 1 unit of *Physarum* M RNase. U+C reaction: 1 μl of RNA was added to 3 μl of buffer III containing 1 unit of *Bacillus cereus* RNase. Alkaline hydrolysis: 1 μl of RNA was added to 2 μl of 50 mM $NaHCO_3$, pH 9.0/1.6 mM EDTA containing 1 μg of carrier tRNA. All reactions were terminated by chilling on ice. For the U+C reaction and the alkaline hydrolysis, 3 μl of loading buffer (0.01% xylene cyanol FF/0.01% bromophenol blue/10 M urea in 0.1× electrode buffer) was added before loading onto the gel. Buffer I: 33 mM Na citrate, pH 5.0/1.7 mM EDTA/0.04% xylene cyanol FF/0.08% bromophenol blue/carrier tRNA at 1 mg/ml/7 M urea. Buffer II: Buffer I, except that the pH of Na citrate was 3.5. Buffer III: as buffer I except that it contained no urea or dyes. Acrylamide gel electrophoresis was as described (4), and the gels were autoradiographed on preflashed film by exposure at −70°C for 1 week with enhancing screens. Shown here is a 20% sequencing gel. Y, ambiguous pyrimidines.

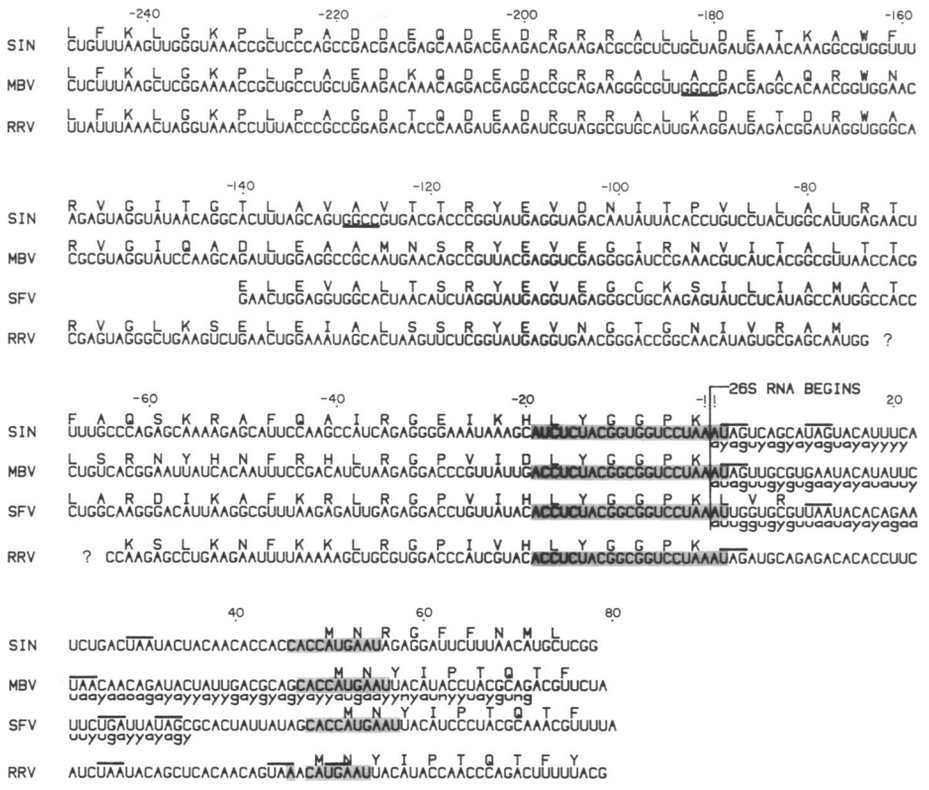

FIG. 2. The nucleotide sequences of the junction regions. The conserved sequences near the start of 26S RNA and around the initiation codon of the capsid proteins are shaded. The deduced protein sequences are shown (one-letter code) above the nucleotide sequences. Nucleotide sequences of the 5′-end of 26S RNA determined by the enzymatic method are shown in lowercase letters. In-phase termination codons for nonstructural proteins are overlined. *Hae* III sites not directly identified are underlined. Numbering of the nucleotides begins with the start of 26S RNA; positive numbers are used for nucleotides in the 26S region and negative numbers, for the 5′-flanking sequence in 49S RNA. Abbreviations: y, pyrimidine; n, any nucleotide; A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; SIN, Sindbis virus; MBV, Middelburg virus; SFV, Semliki Forest virus; RRV, Ross River virus.

G+A reaction to find the junction fragment. A fragment 222 nucleotides long that contained the right purine ladder for the 5′-end sequence of 26S RNA was found and its sequence was then determined (Fig. 2).

A third procedure was used for Ross River virus. RNA was prepared from infected cells; it sedimented at about 26S and contained poly(A). This preparation contained not only 26S RNA but also fragments of degraded 49S RNA and thus the sequences in the junction region were present as well as 26S RNA sequences. cDNA was made with calf thymus primer and the *Hae* III fragments were selected randomly for sequence analysis. The correct fragments were identified by homology. Six nucleotides are apparently missing and we believe them to be present in a small *Hae* III fragment not analyzed. Nonetheless, corresponding sequences can be readily aligned.

The RNA sequences obtained by direct RNA sequence analysis and deduced from cDNA sequences are compared in Fig. 2 and are in agreement. Taken together with the previously published sequence of 26S RNA (5, 12, 13), the complete sequences of both Sindbis virus and Semliki Forest virus 26S RNAs are now known. Sindbis virus 26S RNA is 4,102 nucleotides in length; Semliki Forest virus 26S RNA is 4,074 nucleotides long.

Also shown in Fig. 2 is the deduced amino acid sequences of the COOH-terminal region of the nonstructural proteins (translated from 49S RNA) and the NH_2-terminal region of the capsid protein (translated from 26S RNA). The stop codon for the nonstructural proteins is found at nucleotides 2–4 for three of the viruses and at nucleotides 11–13 for Semliki Forest virus (again, numbering begins with the first nucleotide of 26S RNA). The nonstructural protein sequences are directly compared in Fig. 3. The homology is striking but not uniform, some regions being highly conserved and other regions less well conserved.

Inspection of the junction sequences (Fig. 2) reveals a highly conserved sequence that extends from nucleotides −19 to 2 which we postulate is the recognition signal for the viral transcriptase. A diagrammatic summary of the genomic organization in the junction region is shown in Fig. 4.

DISCUSSION

We have used a comparative approach to study the structure of the junction region of alphavirus 49S RNA. The pronounced homology among the various alphaviruses allows corresponding sequences between any two viruses to be aligned without obtaining the complete RNA sequence of each virus. The homol-

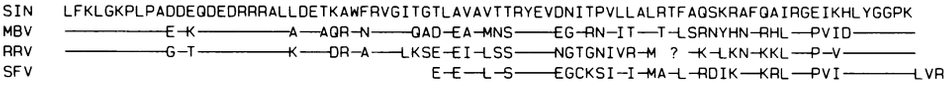

FIG. 3. Comparison of the COOH-terminal sequences of the nonstructural proteins of alphaviruses. The single letter amino acid code virus abbreviations are as in Fig. 2. Dash, amino acid is the same as is found in Sindbis virus.

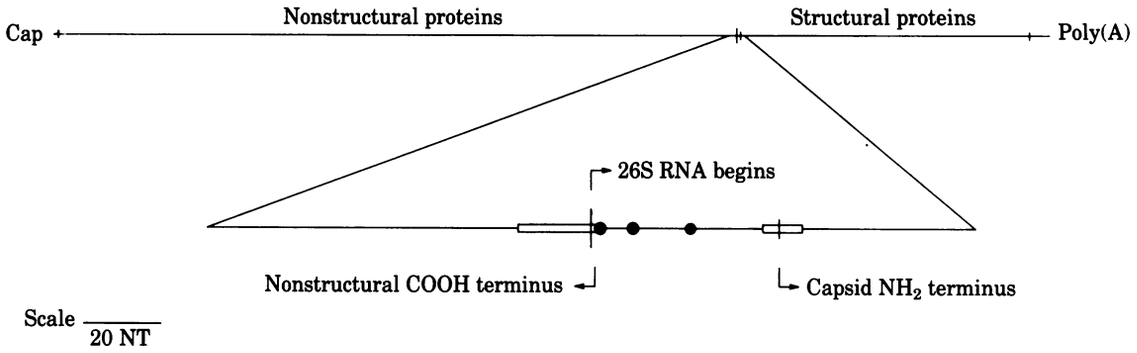

FIG. 4. Schematic of the genome of alphaviruses with the junction region expanded. The boxes represent conserved nucleotide sequences found near the start of 26S RNA and around the start codon of the capsid protein. The start of 26S RNA, the stop point of the nonstructural proteins, and the start of the capsid protein are indicated. In-phase termination codons are shown by solid circles. The figure is drawn to scale for Sindbis virus. The scale bar is for the expanded region.

ogy is much more striking at the protein level than at the RNA level. Use of a computer search program to detect such homology is illustrated in Fig. 5 by a dot matrix in which the deduced COOH-terminal portions of the nonstructual proteins of Sindbis and Middelburg viruses are compared. The comparative approach yields information on the translation phase of the RNA sequences. Inspection of Fig. 2 reveals that, even when the amino acid sequence is conserved, the codons used are not conserved. The evolutionary divergence of these viruses is so extensive that the third codon position in degenerate codon families is essentially randomized. As examples, note that three different codons are used for the Ala at positions −78 to −80 and for the Leu at positions −135 to −137. A second way to illustrate this point is to note that, outside of the conserved sequence near the start of 26S RNA (see below), the Sindbis and Middelburg nonstructural polypeptides shown in Fig. 2 contain the same amino acid at 40 positions, but different codons are used in 25 of these instances. Because of this, translation of corresponding RNA sequences will reveal extensive protein homology only when the proper translation phase is chosen. Thus,

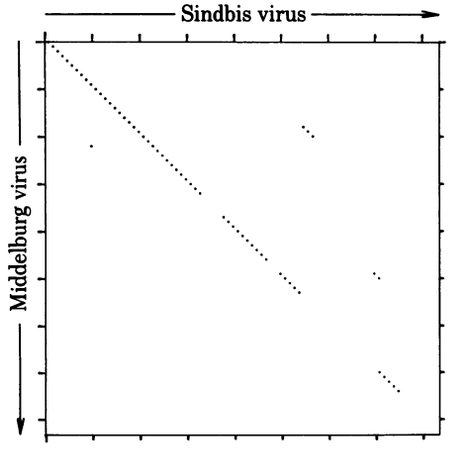

FIG. 5. Dot matrix comparing the amino acid sequence in the COOH-terminal region of the Sindbis virus nonstructural polyprotein to that of Middelburg virus. The amino acid sequences are compared nine residues at a time, and a dot is placed in the matrix whenever at least four amino acids in the two strings are the same and in the same position (i.e., 44% identity of sequence in the string is required). This method has been used to compare nucleotide sequences (15) and has been adapted to comparison of protein sequence. The diagonal line thus created shows the homology between the two proteins. The dots off the diagonal probably result from random matches in the amino acid sequences in these regions. The ends of the arrows mark the COOH termini for the nonstructural proteins.

we are confident that the deduced amino acid sequences shown are correct and that termination of the nonstructural proteins occurs where shown.

It is of interest that the termination codons of the nonstructural proteins lie in the 26S region of the genomic RNA. Thus, the untranslated region between the nonstructural and structural polypeptides is less than 50 nucleotides because the initiation codon of the capsid protein begins at nucleotide 50, 51, 52, and 49 in 26S RNAs of Sindbis, Middelburg, Semliki Forest, and Ross River viruses, respectively. In each case the untranslated region contains a cluster of stop codons that are in phase but slightly separated [three for Sindbis and Semliki Forest, two for Middelburg, four for Ross River (Fig. 2)], implying that multiple stop codons are important. It is also intriguing that Sindbis, Middelburg, and Ross River viruses terminate at the same position, whereas the corresponding Semliki Forest virus codon is a sense codon and a later stop codon (corresponding to the second in-phase Sindbis stop codon) is used which perhaps illustrates the selective advantage of multiple stop codons. A mutation in the first termination codon leading to read through to the second stop codon would lengthen the COOH terminus of the last nonstructural polypeptide which could be important to the virus for evolutionary flexibility in this protein. We note that *in vitro* translation of 49S RNA does not lead to production of structural proteins (reviewed in ref. 3), as would be expected from these results.

A stretch of 21 nucleotides near the beginning of 26S RNA, from nucleotides −19 to 2, is conserved. It is unlikely that conservation of this sequence is caused by a need to conserve the amino acid sequence. As discussed above, conservation of protein sequence does not necessitate such high conservation of the nucleotide sequence. We note that, in contrast to the situation outside this region cited above in which different codons are used for 25 of the 40 conserved amino acids between Sindbis and Middelburg viruses, in this conserved region only 1 of 6 amino acids are encoded by different codons. As another example, we note that the leucine in this region is encoded by CUC in all four viruses, whereas outside this region multiple codons are used for conserved leucine residues (Fig. 2, and see above).

We also have shown that there is a conserved sequence of 19 nucleotides at the 3′ end of the alphavirus genome and have postulated that this forms a replicase recognition site on this plus strand for production of minus strands (4, 6). Similarly, we postulate that the complement of the conserved sequence around the start of 26S RNA forms a recognition site for the viral transcriptase that transcribes 26S RNA from the genomic length minus strand. Because these two recognition sequences are

quite distinct and possess no sequence homology, differential control of 26S RNA and minus strand production is possible. We postulate that different virus polypeptides are required for initiation of minus strand synthesis and for 26S RNA synthesis. These two initiation proteins are probably encoded in different complementation groups and could be components of a viral replicase/transcriptase complex. The nonstructural proteins of Sindbis virus have been grouped into four different complementation groups (16), all of which affect RNA synthesis, and three or four nonstructural polypeptides have been described in infected cells (17).

We suggest that the use of short sequences, 10–20 nucleotides in length, as replication and transcription signals is common among the RNA viruses. The best examples of conserved sequences heretofore have been found in the minus-strand viruses (18). For example, the eight segments of influenza viruses have a highly conserved sequence of 13 nucleotides at both the 3′ and 5′ ends (19). We propose that these form a replication and transcription recognition site and that, because the 3′ and 5′ end sequences are nearly complementary, the same recognition signal is used for both plus-strand and minus-strand synthesis. Similarly, bunyaviruses have a terminal conserved sequence of 9–20 nucleotides, depending on the virus; recombinant viruses containing reassorted genomes can occur only when the genomic segments contain the same or a highly homologous terminal sequence (18), as would be predicted by the hypothesis that they form replication recognition signals in these cases. In other cases it appears to be secondary or tertiary structure in the RNA that is recognized, such as in a number of plant virus RNAs in which specific structures but not sequences are conserved (20). This subject has been reviewed recently in detail (21).

Two stable hairpin structures, each with a ΔG of about -35 kcal (1 cal = 1.84 J) (ΔG calculated as described in ref. 22) can be formed from nucleotides -99 to 67 in the Sindbis virus junction region (unpublished data). It is conceivable that this structure is involved in initiation of 26S RNA synthesis. However, we could not find comparable structures in the three other viruses and, despite its stability, this structure may not possess any particular function in the virus life cycle. This result emphasizes the importance of a comparative approach for identification of sequences and structures of importance in the control of virus replication.

Stable hairpin structures ($\Delta G \approx -12$ kcal) can also be formed in the 5′ untranslated regions of 26S RNA in each of the four viruses, beginning just after the cap, and these may be of importance in translation of the RNA. The size and nucleotide sequence of these structures are not conserved, however.

The tetranucleotide C-A-C-C precedes the AUG initiation codon of the capsid protein in Sindbis, Middelburg, and Semliki Forest viruses, whereas A-A-A-C precedes the initiation codon in Ross River virus. Recently, Kozak (23) surveyed the nucleotides adjacent to the initiation codons from a large number of eukaryotic mRNAs and found that a purine (usually A) is found at position -3 from the initiation AUG and that C is preferred at position -4. The alphavirus 26S RNA initiation sites follow this pattern. However, the nucleotide following the initiation codon, which is usually a G (23), is an A in all of the alphaviruses examined here. These nucleotides are believed to be of importance for selecting the proper AUG codon for initiation of translation, and this idea has been supported by ribosome binding studies *in vitro* (23).

Note Added in Proof. The sequence of the junction region of Semliki Forest virus has been determined recently (24). Our results are in agreement with those reported, with five nucleotide changes, probably due to strain differences, in the region analyzed.

We are grateful to Mrs. E. M. Lenches for preparing chicken embryo cell cultures, to T. Hunkapiller for assistance with the computer work, and to J. R. Bell for critical discussion of the manuscript. This work was supported by Grants PCM 8022830 from the National Science Foundation, GM 06965 and AI 10793 from the National Institutes of Health, and Biomedical Research Support Grant 507RR 07003 from the National Institutes of Health. J.O. was supported in part by Training Grant GM 00086 from the National Institutes of Health; C.M.R. was supported by a fellowship from the Gosney Foundation.

1. Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & VandePol, S. (1982) *Science* **215**, 1577–1585.
2. Webster, R. G., Laver, W. G., Air, G. M. & Schild, G. C. (1982) *Nature (London)* **296**, 115–121.
3. Strauss, J. H. & Strauss, E. G. (1977) in *The Molecular Biology of Animal Viruses*, ed. Nayak, D. P. (Dekker, New York), pp. 111–166.
4. Ou, J.-H., Strauss, E. G. & Strauss, J. H. (1981) *Virology* **109**, 281–289.
5. Rice, C. M. & Strauss, J. H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2062–2066.
6. Ou, J.-H., Trent, D. & Strauss, J. H. (1982) *J. Mol. Biol.* **156**, 719–730.
7. Monroe, S. S., Ou, J.-H., Rice, C. M., Schlesinger, S., Strauss, E. G. & Strauss, J. H. (1981) *J. Virol.* **41**, 153–162.
8. Paoletti, E., Rosemond-Hornbeak, H. & Moss, B. (1974) *J. Biol. Chem.* **249**, 3273–3280.
9. Rose, J. K. & Lodish, H. F. (1976) *Nature (London)* **262**, 32–37.
10. Ahlquist, P., Dasgupta, R., Shih, D. S., Zimmern, D. & Kaesberg, P. (1979) *Nature (London)* **281**, 277–282.
11. Rice, C. M. & Strauss, J. H. (1981) *J. Mol. Biol.* **150**, 315–340.
12. Garoff, H., Frischauf, A.-M., Simons, K., Lehrach, H. & Delius, H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6376–6380.
13. Garoff, H., Frischauf, A.-M., Simons, K., Lehrach, H. & Delius, H. (1980) *Nature (London)* **288**, 236–241.
14. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
15. Hieter, P. A., Max, E. E., Seidman, J. G., Maizel, J. V. & Leder, P. (1980) *Cell* **22**, 197–207.
16. Strauss, E. G. & Strauss, J. H. (1980) in *The Togaviruses*, ed. Schlesinger, R. W. (Academic, New York), pp. 393–426.
17. Schlesinger, M. J. & Kääriäinen, L. (1980) in *The Togaviruses*, ed. Schlesinger, R. W. (Academic, New York), pp. 371–392.
18. Bishop, D. H. L., Clerx, J. P. M., Clerx-Van Haaster, C. M., Robeson, G., Rozhon, E. J., Ushijima, H. & Veerisetty, V. (1981) in *The Replication of Negative Strand Viruses*, eds. Bishop, D. H. L. & Compans, R. W. (Elsevier/North-Holland, New York), pp. 135–146.
19. Robertson, J. S. (1979) *Nucleic Acids Res.* **6**, 3745–3757.
20. Ahlquist, P., Dasgupta, R. & Kaesberg, P. (1981) *Cell* **23**, 183–189.
21. Strauss, E. G. & Strauss, J. H. (1982) *Curr. Top. Microbiol. Immunol.*, in press.
22. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London)* **246**, 40–41.
23. Kozak, M. (1981) *Nucleic Acids Res.* **9**, 5233–5252.
24. Riedel, H., Lehrach, H. & Garoff, H. (1982) *J. Virol.* **42**, 725–729.