Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 5271–5275, September 1983
Biochemistry

Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon

(Sindbis virus/nucleotide sequence/Middelburg virus)

ELLEN G. STRAUSS, CHARLES M. RICE, AND JAMES H. STRAUSS

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

Communicated by Ray D. Owen, May 31, 1983

ABSTRACT **We have obtained the nucleotide sequence of the genomic RNAs of two alphaviruses, Sindbis virus and Middelburg virus, over an extensive region encoding the nonstructural (replicase) proteins. In both viruses in an equivalent position an opal (UGA) termination codon punctuates a long otherwise open reading frame. The nonstructural proteins are translated as polyprotein precursors that are processed by posttranslational cleavage into four polypeptide chains; the sequence data presented here indicate that the COOH-terminal polypeptide, ns72, may be produced by read-through of this opal codon. The high degree of amino acid homology between the ns72 polypeptides of the two viruses, in contrast to the lack of conserved sequence upstream from the read-through site, suggests that ns72 plays an important role in viral replication, possibly modulating the action of other replicase components.**

The alphavirus genus of the family Togaviridae contains about 20 members, many of which are important human or veterinary pathogens. These viruses contain an infectious RNA of about 11,700 nucleotides, which sediments at 49 S (reviewed in ref. 1). In the virus this RNA is surrounded by a single species of basic protein, the capsid protein, to form a nucleocapsid with cubic symmetry. As the final event in maturation this capsid buds through the plasma membrane of the infected cell and acquires an envelope containing lipids of host cell origin and two virus-encoded transmembrane glycoproteins, termed E2 and E1.

The alphaviruses produce two messenger RNAs after infection (1). The first, 49S RNA, appears to be identical to the RNA in the virion and is translated as polyproteins that are processed to form the nonstructural proteins (2) required to replicate the virus RNA and to transcribe a subgenomic 26S RNA. This 26S RNA is the major messenger found on host cell polysomes and is translated into the structural proteins found in virions. The sequences of the 26 RNAs of three alphaviruses, Semliki Forest virus (3, 4), Sindbis virus (5), and Ross River virus (6), have been determined recently and the details of the encoded proteins established. We have previously reported that the translation of the nonstructural polyproteins from alphavirus 49S RNA terminates with multiple in-phase stop codons positioned about two-thirds of the distance from the 5′ end (7). In this communication we report that translation of the nonstructural proteins is modulated by the presence of an opal codon within the RNA.

MATERIALS AND METHODS

Preparation of RNA. Growth of Sindbis virus or Middelburg virus and preparation of 49S RNA from infected cells or from purified virions have been described (8).

Sequence of Sindbis RNA. Sindbis 49S RNA was transcribed into DNA by using reverse transcriptase and random priming as described (9). The single-stranded (ss) cDNA was cut with restriction enzymes *Hae* III or *Taq* I and the resulting fragments were subjected to sequence analysis by the chemical methods of Maxam and Gilbert (10), as described (9). All but one of the *Hae* III fragments longer than 30 nucleotides and about 30 *Taq* I fragments that contained non-26S sequence were subjected to sequence analysis. At this stage, together with the sequence of the 26S RNA (5), about 90% of the 49S sequence had been obtained. These sequences were aligned by restriction enzyme analysis of double-stranded (ds) cDNA made to 49S RNA by using infrequently cutting enzymes such as *Eco*RI (6). At this point fragments encompassing the entire 49S sequence could be aligned but the sequence still possessed gaps and ambiguities due to compression artifacts. Using this information to devise a cloning strategy, we obtained the entire genome as four nonoverlapping clones in a plasmid derivative of pBR322 (unpublished data). Cloned DNA was used to sequence selected regions of the genome by the chemical method. A detailed description of the methods used and the entire sequence of the Sindbis 49S RNA will be presented elsewhere.

Sequence Analysis of Middelburg RNA. A large number of *Hae* III fragments from ss cDNA made to Middelburg 49S RNA were subjected to sequence analysis as previously described (9). These sequences were translated in all three reading frames and the deduced amino acid sequences were compared to those of Sindbis virus by using computer homology routines. In this way most of the sequences of the Middelburg fragments obtained could be positioned relative to the Sindbis 49S RNA sequence and a partial restriction map was generated by computer. Long ds cDNA was cut with infrequently cutting enzymes and the resulting fragments were subjected to sequence analysis to confirm the alignment and obtain additional sequence. Only two *Eco*RI sites were found to be present in Middelburg RNA, bracketing the region containing the stop codon. ds Middelburg cDNA was cut with *Eco*RI and a 2.5-kilobase fragment was cloned into the *Eco*RI site of a plasmid derived originally from pBR322. This cloned fragment was subjected to sequence analysis in its entirety. A sequence analysis schematic for both viruses is shown in Fig. 1.

RESULTS AND DISCUSSION

During sequence analysis of Sindbis 49S RNA we discovered the presence of an opal stop codon within the region encoding the nonstructural proteins. A simple map of the Sindbis genome (Fig. 2) illustrates the location of this opal codon in relation to the regions encoding the nonstructural proteins as well as the structural polypeptides of the virion. To demonstrate the generality of an open reading frame for the nonstructural pro-

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.

Abbreviations: ss, single-stranded; ds, double-stranded.

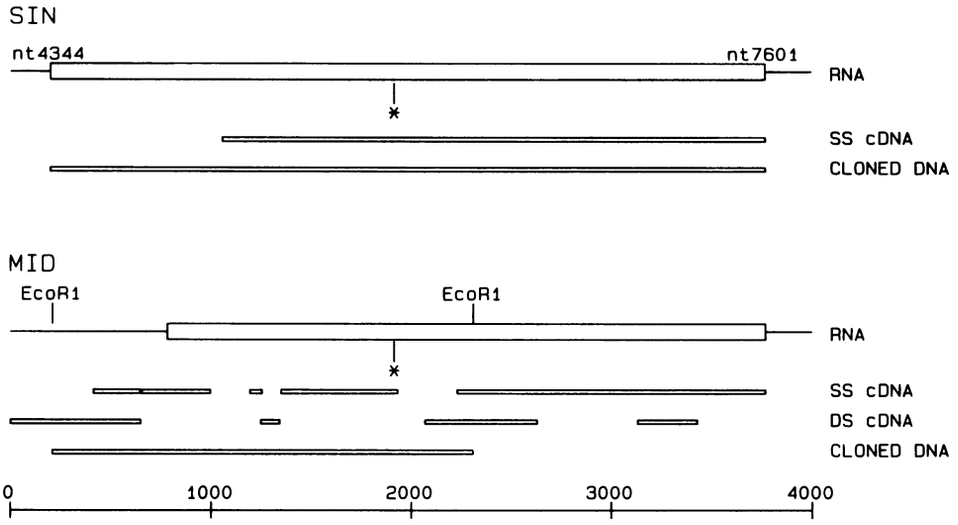

FIG. 1. Sequence analysis schematic for Sindbis virus (SIN) RNA and Middelburg virus (MID) RNA. The Sindbis sequence is numbered from the 5′ terminus of 49S RNA. The two *Eco*RI sites define the cloned segment of the Middelburg genome. The regions sequenced as ss cDNA, as ds cDNA, and as cloned DNA, respectively, are noted for each virus. The open boxes are the regions of sequence presented in Fig. 3; the asterisk marks the opal codon. nt, Nucleotide.

teins punctuated with a termination codon as an integral mechanism of alphavirus replication, we obtained the nucleotide sequence of a second alphavirus, Middelburg virus, in the region shown by a solid box in Fig. 2. The translated sequences for both Sindbis and Middelburg viruses in this region are shown in Fig. 3. In the case of Sindbis virus, nucleotides are numbered from the 5′ terminus of the RNA and amino acids from the NH_2 terminus of the nonstructural polyprotein precursor. The sequence of Middelburg is not complete and its sequence is aligned by homology with the corresponding Sindbis sequence. Both viruses have an opal codon at the same position in the sequence (boxed in Fig. 3). For Sindbis virus this stop codon interrupts an otherwise open reading frame for 2,513 amino acids and is found at position 1,897 in this sequence (unpublished data). For both viruses the sequence through the stop codon was obtained with both ss cDNA as well as cloned DNA (Fig. 1). As noted earlier, the sequence of ss cDNA gives the majority nucleotide at any position for the population of reverse transcriptase copies produced from the population of RNA molecules and thus cloning artifacts, reverse transcriptase errors, or minor components in the RNA population do not influence the results. Furthermore, in the case of Sindbis virus both strands of the cloned cDNA were subjected to sequence analysis in both directions (i.e., both 5′ to 3′ and 3′ to 5′) to rule out sequence analysis artifacts. The existence of the identical stop codon in the same position in two different alphaviruses makes it clear that this opal codon is a fundamental landmark in the alphavirus genome.

Downstream of the terminator, there is a long open reading frame encoding highly conserved proteins in the two viruses (Fig. 3 and also see below), indicating that this sequence must be translated *in vivo* to produce a protein important for virus replication. Direct evidence for the existence of such a polypeptide, called ns72, comes from experiments with an antibody produced in rabbits to the synthetic polypeptide Arg-Gly-Glu-Ile-Lys-His-Leu-Tyr-Gly-Gly-Pro-Lys, which is the deduced amino acid sequence at the COOH terminus of this polypeptide for Sindbis virus (overlined in Fig. 3). This antipeptide antibody precipitates ns72 as well as several larger polypeptides, presumably precursors, from Sindbis-infected cells (S. Lopez and J. Bell, personal communication). Polypeptide ns72 could be produced by several mechanisms, including translation from a spliced message, internal initiation, or read-through synthesis. The presence of a minor spliced message seems unlikely in view of the fact that spliced mRNAs have not been detected for animal RNA viruses whose messages are produced solely in the cytoplasm and no evidence for such a RNA has been found for alphaviruses (12). Similarly, no indication of internal initiation has been seen for alphaviruses either *in vivo* (2) or *in vitro* (13). Moreover, the first methionine residue downstream of the opal codon is located at residue 1,957 and would result in a polypeptide only 556 amino acids long. Because the open reading frames upstream and downstream of the termination signal are in phase, the most likely mechanism for producing ns72 that fits all of the existing data is read-through synthesis to produce a full-length polyprotein precursor of 2,513 amino acids, followed by posttranslational proteolytic cleavage. Indeed, minor amounts of such a precursor have been detected during *in vitro* translation of 49S RNA by Collins *et al.* (13), who suggested on the basis of the synthesis kinetics that this polypeptide might be produced by read-through of a stop codon. Moreover, the largest polypeptide precipitated from infected cells by the antibody to the COOH-terminal peptide is sufficiently large to be the 2,513 amino acid protein encoded by the entire nonstructural region (S. Lopez and J. Bell, personal communication).

Thus, we believe that translation of Sindbis virus 49S RNA produces two polyproteins, which contain the nonstructural proteins, both of which are initiated at the same location near

FIG. 2. Map of the Sindbis virus genome. Boxes indicate translated regions. The solid box corresponds to the open box in Fig. 1 and to the sequences shown in Fig. 3; the asterisk marks the opal codon. The crosshatched box is the region translated from the 26S subgenomic RNA into the structural proteins.

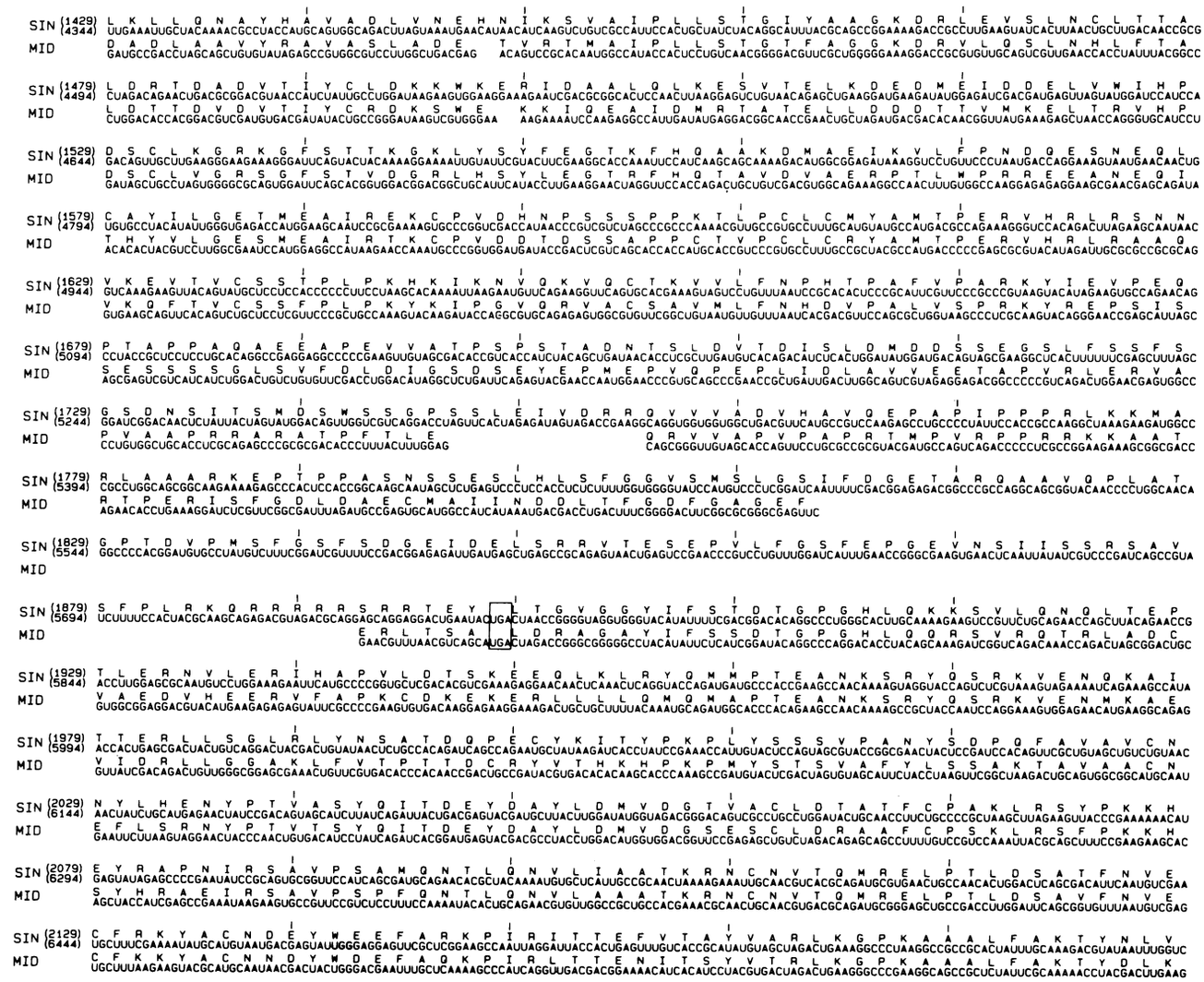

FIG. 3. Translated nucleotide sequence of Sindbis virus (SIN) and Middelburg virus (MID) RNAs in the region containing the opal stop codon. For Sindbis virus nucleotides are numbered from the 5′ terminus of 49S RNA and amino acids are numbered from the NH_2 terminus (11) of the nonstructural polyprotein precursor. The Middelburg sequence is aligned by homology with the Sindbis sequence and the two underlined *Hae* III sites (G-G-C-C) were not directly sequenced; the positioning of the Middelburg sequence between nucleotides 5,394 and 5,729 of Sindbis is arbitrary because of the lack of homology and the large deletion in the Middelburg sequence relative to the Sindbis sequence. In other regions gaps have been introduced as necessary to maintain the homologous alignment. The opal stop codons are boxed; the wavy overline indicates the peptide synthesized to produce antibodies. The single-letter amino acid code is used in the translation of the RNA and the amino acid is positioned above the first nucleotide of the corresponding codon. 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, and Y = Tyr.

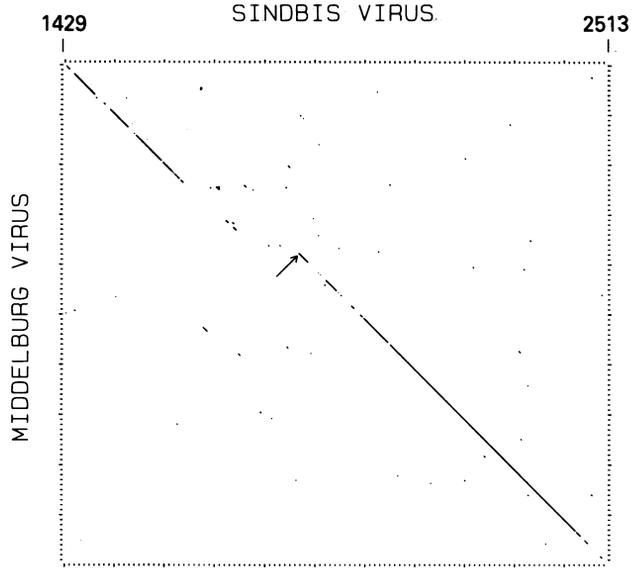

FIG. 4. Dot matrix comparing the proteins encoded by Sindbis RNA and Middelburg RNA around the opal termination codon. The protein sequences are compared nine amino acids at a time and a dot is placed in the matrix whenever five match. The position of the opal termination codon is indicated by the arrow.

the 5′ end. One precursor is 1,896 amino acids long, terminates at the opal codon, and is cleaved to produce ns60, ns89, and ns76. The second precursor is produced in relatively smaller amounts by read-through of the opal codon, is 2,513 amino acids long, terminates at the multiple in-phase termination codons in the junction region, and is cleaved to produce a fourth protein, ns72, presumably in addition to the three products mentioned above (Fig. 2). ns60, ns89, and ns76 have been observed in both lysates of infected cells (reviewed in ref. 2) and as the products of *in vitro* translation of 49S RNA (13) but the small amounts of ns72 produced have only been visualized by immunoprecipitation.

The protein-encoding regions on either side of the opal termination codon have several interesting features, which are illustrated in Fig. 4. Fig. 4 shows a dot matrix in which the amino acid sequences of Sindbis virus throughout the region contained in Fig. 3 are compared with those of Middelburg virus nine residues at a time. A dot is placed in the matrix whenever five residues in the string of nine match. The diagonal lines show regions of protein homology between the two viruses; off-diagonal marks are due to random matches. It is obvious that the two viruses are fairly closely related and Fig. 4 also illustrates the principle used for alignment of Middelburg sequence data with that of Sindbis virus. Downstream from the opal termination codon (arrow in Fig. 4) the two sequences are highly homologous and upstream from this stop codon, between amino acids 1,429 and 1,676 in the Sindbis sequence, the two sequences also demonstrate considerable homology. However, between amino acids 1,676 and 1,896 in the Sindbis sequence there is virtually no detectable homology. Furthermore, as the shift in the diagonal makes clear, there is a discontinuity in the alignment and the Sindbis sequence is longer in this region by 88 amino acids. Thus, the COOH-terminal sequence of the polyprotein preceding the opal termination codon is not homologous in these two alphaviruses and differs in length so that the third nonstructural protein of Sindbis virus, ns76 (Figs. 2 and 3), is larger than the corresponding Middelburg protein. The function of this region of protein is unknown but the variation in size and sequence might imply that it is not vital for the virus replication cycle.

In contrast, the extreme COOH-terminal nonstructural protein, ns72, is highly conserved between Middelburg and Sindbis (Figs. 4 and 5). Beginning just before the opal codon and continuing to the series of in-phase stop codons in the junction region (7), which begin with an amber codon corresponding to nucleotides 2–4 of 26S RNA, the two sequences are identical in length, exhibit perfect alignment with no gaps or insertions, and show an overall homology of 73%. The start of ns72 is unclear but estimates of its molecular weight based on migration in acrylamide gels and the fact that there is no conservation in length or sequence upstream from the opal codon makes it likely that ns72 begins very near this stop codon, a likely start being the glycine six residues downstream (unpublished observations). If we assume they begin at the stop codon, then the two ns72s have the same amino acid at 448 out of 616 positions. In addition, 33 out of the 168 substitutions that occur are conservative. There is one stretch of 98 amino acids (amino acids 2,161–2,259 in the Sindbis sequence) in which there are only five differences (Fig. 5). The COOH-terminal regions show less

```
SIN (1429) LKLLQNAYHAVADLVNEMNIKSVAIPLLSTGIYAAGKDRLEVSLNCLTTALDRTDADVTIYCLDKKWKERIDAALQLKESVTELKDEDMEIDDELVWIHP
MID        DAD.AAV.R...S.AD. TVRTM.......TF.G....VLQ...H.F....T..V......R..S.E KKIQEAIDMRTA...L.D.TTVMK..TRV..

SIN (1529) DSCLKGRKGFSTTKGKLYSYFEGTKFHQAAKDMAEIKVLFPNDQESNEQLCAYILGETMEAIREKCPVDHNPSSSPPKTLPCLCMYAMTPERVHRLRSNN
MID        ....V..S....VD.R.H..L...R...T.V.V..RPT.W.RRE.A...ITH.V...S.....T.....DTD..A..C.V....R.............AAQ

SIN (1629) VKEVTVCSSTPLPKHKIKNVQKVQCTKVVLFNPHTPAFVPARKYIEVPEQPTAPPAQAEEAPEVVATPSPSTADNTSLDVTDISLDMDDSSEGSLFSSFS
MID        ..QF.....F....Y..PG..R.A.SA.M...HDV..L.SP...R.PSISSESSSSGLSVFDLDIGSD.EYEPMEPVQPEPL.D.AVVEETAPVRLERVA

SIN (1729) GSDNSITSMDSWSSGPSSLEIVDRQVVVADVHAVQEPAPIPPPRLKKMARLAAARKEPTPPASNSSESLHLSFGGVSMSLGSIFDGETARQAAVQPLAT
MID        PVAAPRRARATPFTLE          .R...P.P.PRTMPVR..R.K.AAT.TPERISFGDLD.ECMAIINDDLTF.DFGAGEF

SIN (1829) GPTDVPMSFGSFSDGEIDELSRRVTESEPVLFGSFEPGEVNSIISSRSAVSFPLRKQRRRRRSRRTEYXLTGVGGYIFSTDTGPGHLQKKSVLQNQLTEP
MID                                                                    E.L.SA..DRA.A....S.........QR..R.TR.ADC

SIN (1929) TLERNVLERIHAPVLDTSKEEQLKLRYQMMPTEANKSRYQSRKVENQKAITTERLLSGLRLYNSATDQPECYKITYPKPLYSSSVPANYSDPQFAVAVCN
MID        VA.DVHE..VF..KC.KE..RL.L.QM..A.............M..EVID...G.AK.FVTP.TDCRYVTHKH...M..T...AFYL.SAKT...A..

SIN (2029) NYLHENYPTVASYQITDEYDAYLDMVDGTVACLDTATFCPAKLRSYPKKHEYRAPNIRSAVPSAMQNTLQNVLIAATKRNCNVTQMRELPTLDSATFNVE
MID        EF.SR.....T..............SES...R.A...S....F....S.HRAE.......PF.........A.....................V....

SIN (2129) CFRKYACNDEYWEEFARKPIRITTEFVTAYVARLKGPKAAALFAKTYNLVPLQEVPMDRFVMDMKRDVKVTPGTKHTEERPKVQVIQAAEPLATAYLCGI
MID        ..K.....ND..D...Q....L...NI.S..T.............D.K...........V.....................................

SIN (2229) HRELVRRLTAVLLPNIHTLFDMSAEDFDAIIAEHFKQGDPVLETDIASFDKSQDDAMALTGLMILEDLGVDQPLLDLIECAFGEISSTHLPTGTRFKFGA
MID        ........N.....V..............S...RP..A..............SL.Y....L...........E...AS....T..............

SIN (2329) MMKSGMFLTLFVNTVLNVVIASRVLEERLKTSRCAAFIGDDNIIHGVVSDKEMAERCATWLNMEVKIIDAVIGERPPYFCGGFILQDSVTSTACRVADPL
MID        ..............M..MT..........TN.K..........V...K...LL.....A.M..........MC............VF.Q..G.C.......

SIN (2429) KRLFKLGKPLPADDEQDEDRRRALLDETKAWFRVGITGTLAVAVTTRYEVDNITPVLLALRTFAQSKRAFQAIRGEIKHLYGGPK
MID        ............E.K.........A..AQR.N....QAD.EA.MNS....EG.RN.IT..T.LSRNYHN.RHL..PVID......
```

FIG. 5. Direct comparison of the amino acid sequences encoded by Sindbis RNA and Middelburg RNA around the opal termination codon. The sequence of Sindbis is given in the first line. Amino acids are numbered starting with the NH_2 terminus of the polyprotein precursor. The Middelburg sequence is shown below; a dot means the amino acid is the same as in the Sindbis sequence. Gaps have been introduced in the Middelburg sequence for the purpose of alignment. The single-letter code is used as in Fig. 3; the boxed X is the position of the opal codon.

homology (Fig. 5). We have postulated that components of the alphavirus replicase recognize specific nucleotide sequences of about 20 nucleotides in length in the RNA as promoter sites (7, 11, 14) and it is tempting to speculate that the highly conserved core of ns72 has such a function. In any event, the high degree of homology between these two viruses indicates an important function for ns72 in the virus replication cycle. As has been previously noted (7), the alphaviruses have diverged sufficiently such that the codon used for any particular amino acid has been randomized, even in highly homologous segments. In ns72 different codons are used between Sindbis and Middelburg for the same amino acid in 57% of the cases in which conservation occurs (Fig. 3). The slight difference from complete randomization (66% difference) probably arises because some codons are infrequently used by alphaviruses (5, 6). Thus, conservation of amino acid sequence is highly significant and the translation frame can be deduced by comparison of conserved amino acid sequences; because of the divergence in codon usage the highest degree of protein sequence homology is observed only when the proper translation frames are being compared.

Although termination codons have been located within the translated regions of a number of viruses, at the current time the alphaviruses appear unique among animal viruses in using read-through to modulate the synthesis of their proteins. An amber codon punctuates the genome of the murine retroviruses between the *gag* and *pol* genes (15) and it was originally believed that read-through was the mechanism for producing relatively minor amounts of the polymerase products (16, 17). However, recent results with an avian retrovirus, Rous sarcoma virus, have revealed that in this case the downstream, open reading frame for *pol* is out of phase with the *gag* gene (18) and thus the polymerase moieties may be produced from a minor species of messenger produced by splicing in which the amber codon could have been deleted. Amber codons are also present in the genomes of a variety of RNA plant viruses [including tobamoviruses (19) and tymoviruses (20)] and read-through products have been obtained by translation of these and other plant virus RNAs in *in vitro* protein synthesis systems (20–24). Whether these experiments accurately reflect the conditions of replication during normal infection is unclear at this time.

The mechanism by which read-through of an opal codon might occur during alphavirus infection is unknown. Specific suppression of the opal codon is unlikely because in the case of Sindbis virus the translation of the structural proteins is terminated with an opal codon (5). The simplest mechanism predicts that read-through is a probablistic event of relatively low frequency, perhaps due to wobble in the first anticodon position to insert a tryptophan. If so, accumulation of ns72 would be relatively slow in the infected cell and functionally ns72 would appear as a late protein. An attractive hypothesis for the function of such a polypeptide is as a controlling or modifying factor for the polymerase, perhaps to effect the shutoff of minus-strand template synthesis that occurs at 3–4 hr postinfection (25, 26).

Thus, alphaviruses, with a small and relatively simple genome, are capable of regulating the synthesis of their virus-encoded products at three different levels by two different mechanisms. Production of a subgenomic (26S) RNA as the predominant message in infected cells insures that most of the virus-specific protein synthesis is devoted to production of the structural proteins. Translation of the genomic RNA up to the opal codon produces three nonstructural polypeptides in lesser amounts and read-through may produce a fourth polymerase component in very small quantities for specific regulation of replication.

The computer programs used in this work, including production of figures, were written by Tim Hunkapiller and we are grateful to him for access to these programs and instructions in their use. The computer work was performed on the computer facility of Lee Hood and we are grateful for the time on these facilities. Edith Lenches gave expert technical assistance in the maintenance and preparation of virus stocks and cell lines. This work was supported by Grants AI 10793 and GM 06965 from the National Institutes of Health and by Grant PCM 8022830 from the National Science Foundation.

1. Strauss, J. H. & Strauss, E. G. (1977) in *The Molecular Biology of Animal Viruses*, ed. Nayak, D. P. (Dekker, New York), pp. 111–166.
2. Schlesinger, M. J. & Kääriäinen, L. (1980) in *The Togaviruses*, ed. Schlesinger, R. W. (Academic, New York), pp. 371–392.
3. Garoff, H., Frischauf, A.-M., Simons, K., Lehrach, H. & Delius, H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6376–6380.
4. Garoff, H., Frischauf, A.-M., Simons, K., Lehrach, H. & Delius, H. (1980) *Nature (London)* **288**, 236–241.
5. Rice, C. M. & Strauss, J. H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2062–2066.
6. Dalgarno, L., Rice, C. M. & Strauss, J. H. (1983) *Virology*, in press.
7. Ou, J.-H., Rice, C. M., Dalgarno, L., Strauss, E. G. & Strauss, J. H. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5235–5239.
8. Ou, J.-H., Strauss, E. G. & Strauss, J. H. (1981) *Virology* **109**, 281–289.
9. Rice, C. M. & Strauss, J. H. (1981) *J. Mol. Biol.* **150**, 315–340.
10. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
11. Ou, J.-H., Strauss, E. G. & Strauss, J. H. (1983) *J. Mol. Biol.*, in press.
12. Strauss, E. G. & Strauss, J. H. (1983) *Curr. Top. Microbiol. Immunol.* **105**, 1–97.
13. Collins, P. L., Fuller, F. J., Marcus, P. I., Hightower, L. E. & Ball, L. A. (1982) *Virology* **118**, 363–379.
14. Ou, J.-H., Trent, D. G. & Strauss, J. H. (1982) *J. Mol. Biol.* **156**, 719–730.
15. Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. (1981) *Nature (London)* **293**, 543–548.
16. Philipson, L., Anderson, P., Olshevsky, U., Weinberg, R., Baltimore, D. & Gesteland, D. (1978) *Cell* **13**, 189–199.
17. Oppermann, H., Bishop, J. M., Varmus, H. E. & Levintow, L. (1977) *Cell* **12**, 993–1005.
18. Schwartz, D., Tizard, R. & Gilbert, W. (1983) *Cell* **32**, 853–869.
19. Goelet, P., Lomonossoff, G. P., Butler, P. J. G., Akam, M. E., Gait, M. J. & Karn, J. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5818–5822.
20. Morch, M.-D., Drugeon, G. & Benicourt, C. (1982) *Virology* **119**, 193–198.
21. Pelham, H. R. B. (1978) *Nature (London)* **272**, 469–471.
22. Mang, K.-G., Ghosh, A. & Kaesberg, P. (1982) *Virology* **116**, 264–274.
23. Dougherty, W. G. & Kaesberg, P. (1981) *Virology* **115**, 45–56.
24. Bisaro, D. M. & Siegel, A. (1982) *Virology* **118**, 411–418.
25. Sawicki, D. L. & Sawicki, S. G. (1980) *J. Virol.* **34**, 108–118.
26. Sawicki, D. L., Sawicki, S. G., Keränen, S. & Kääriäinen, L. (1981) *J. Virol.* **39**, 348–358.