JOURNAL OF VIROLOGY, Aug. 1989, p. 3345–3352
0022-538X/89/083345-08$02.00/0

Vol. 63, No. 8

In Vitro Processing of Dengue Virus Structural Proteins: Cleavage of the Pre-Membrane Protein

LEWIS MARKOFF

Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 15 February 1989/Accepted 19 April 1989

Processing of dengue virus structural proteins was assessed in vitro. RNA transcripts for cell-free translation were prepared from cloned DNA (dengue virus type 4, strain 814669 genome) encoding capsid, pre-membrane (prM), and the first 23 amino acids of envelope (E). Processing of a 33-kilodalton precursor polypeptide encoded by wild-type RNA transcripts occurred only in the presence of added microsomal membranes. Under these conditions, cleavage at the capsid-prM and prM-E sites and glycosylation of prM occurred in association with translocation. Amino acid sequence analysis confirmed that translation initiated at the predicted N terminus of the capsid and that capsid-prM cleavage occurred at the predicted site for the action of signal peptidase following a candidate signal sequence (hydrophobic residues 100 to 113) in the dengue virus precursor. Mutations were introduced into the dengue virus DNA template by site-directed mutagenesis, altering nucleotide sequences encoding the capsid and the candidate signal for prM. The phenotypes of the mutants were deduced by analysis of the products of cell-free translation of the respective RNA transcripts. The resulting observations confirmed that cleavage at the capsid-prM and prM-E sites is effected entirely by signal peptidase and that the candidate signal is required for translocation.

The four serotypes of dengue virus are members of the flavivirus family, which now includes more than 70 different viruses. Flavivirus particles are composed of three structural proteins: a genome-associated capsid protein, a membrane-associated (M) protein that is derived during virus maturation by internal cleavage from a glycosylated precursor protein (prM), and a membrane-anchored hemagglutinating envelope glycoprotein (E). E is the major antigenic determinant of free virus responsible for serotype specificity (18–20, 31, 32).

Flavivirus genomic RNAs are 10,500 to 11,000 nucleotides long and include a single open reading frame encoding a precursor polyprotein approximately 3,400 amino acids long (5–7, 10, 14–16, 18, 22, 26, 33). Structural proteins are derived by cleavage from the N terminus of the precursor in the following order: capsid, prM, and E (2–4, 18, 22, 24, 28). Since some viral proteins are glycosylated and virus particles are assembled in the endoplasmic reticulum (11, 19), processing of the flavivirus precursor polyprotein must include not only cleavage but also translocation of cleavage products. An association between cleavage of capsid and E proteins and cotranslational translocation has been suggested by the results of an in vitro study (25). In addition, the dengue virus capsid, prM, E, and NS1 were apparently normally cleaved and glycosylated when expressed in a vaccinia virus recombinant containing a 4-kilobase (kb) fragment of cloned dengue virus type 4 DNA encoding the capsid, prM, E, NS1, and NS2a (34). Therefore, nonstructural proteins (NS2b through NS5 [14]) not encoded in the vaccinia virus recombinant are not required to process the structural proteins and NS1.

As expected, the locations of potential signal and transmembrane segments in the flavivirus precursor are conserved. Within N-terminal sequences encoding structural proteins, the precursor contains an uninterrupted 14-amino-acid hydrophobic segment (residues 100 to 113 in the precursor sequence). This sequence precedes the N terminus of prM determined by amino acid sequencing of labeled proteins isolated from virus-infected cells (2, 4) and is the candidate signal sequence for translocation of prM (18, 19, 33). The candidate signal is preceded by a conserved pair or triplet of basic amino acids (Arg-Lys-Arg; residues 97 to 99 in the dengue virus type 4 sequence [33]). It has been suggested that these conserved residues might constitute a proteolytic site (18, 33). In addition, the predicted C termini of both prM and E are hydrophobic segments (residues 246 to 279 and 735 to 773, respectively) asymmetrically interrupted in each case by a single Arg. These are candidate transmembrane segments for prM and E and signals for translocation of the respective downstream glycoproteins, E and NS1 (14, 18, 22, 33).

To establish the association between cleavage of structural proteins from the precursor and translocation, processing of a truncated dengue virus precursor protein was studied in vitro. RNA transcripts were prepared from a 911-base-pair fragment of dengue virus type 4 DNA, including the complete capsid and prM nucleotide sequences and sequences encoding the first 23 amino acids of E. Cleavage of the encoded precursor at the capsid-prM and prM-E sites was dependent upon the presence of microsomes during cell-free translation and appeared to mimic the in vivo events. To determine the mechanism of cleavage, mutations of nucleotide sequences encoding the capsid and the candidate prM signal were introduced into dengue virus DNA by site-directed mutagenesis. The results of cell-free translation of mutant RNA transcripts are also described. An intact signal at residues 100 to 113 was required for translocation and cleavage of both prM and E. Mutations of the capsid sequence had no effect on processing, and cleavage appeared to be entirely mediated by the signal peptidase.

MATERIALS AND METHODS

Preparation of recombinant pTZ/Sal. A 4,040-base-pair (4-kb) fragment of cloned dengue virus type 4 strain 814669 DNA bounded by *Bgl*II restriction endonuclease cleavage sites at positions 88 and 4128 was subcloned into the plasmid vector pTZ18U (Pharmacia, Inc., Piscataway, N.J.) at the *Bam*HI site within the pTZ18U polylinker in plus-sense

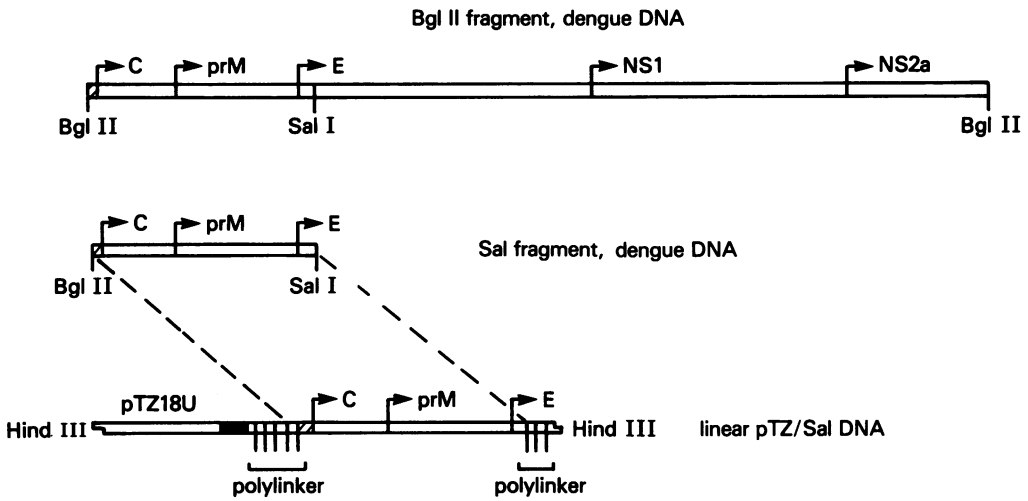

FIG. 1. Derivation of *Sal* DNA in the pTZ/Sal construct. A 4-kb fragment of cloned dengue virus (type 4, strain 814669) DNA previously expressed in a vaccinia virus recombinant (34) was cloned into the plasmid vector PTZ18U at the *Bam*HI site. pTZ18U-dengue virus recombinant DNA was digested with *Sal*I, excising 3′-terminal sequences from the dengue virus DNA. The resultant pTZ/Sal recombinant DNA linearized by *Hin*dIII digestion, used as the template for synthesis of RNA transcripts, is depicted. ▨, Noncoding sequences in dengue virus DNA; ■, T7 RNA polymerase promoter in pTZ18U.

relationship to the T7 RNA polymerase promoter sequence contained in the vector. 3′-Terminal nucleotide sequences in the dengue virus DNA were excised by digestion of the recombinant vector DNA with the restriction endonuclease *Sal*I, which cleaves the dengue virus DNA at nucleotide 999 and pTZ18U DNA at a unique *Sal*I site in the polylinker 3′ to dengue virus DNA sequences. The *Sal*I-digested DNA was religated to construct pTZ18U-dengue virus recombinant DNA which included only the 5′ 911 nucleotides (88 to 999) of the 4-kb dengue virus DNA fragment. These 911 nucleotides define the *Sal* fragment of dengue virus DNA. pTZ18U recombinant DNA containing the *Sal* fragment of dengue virus DNA is designated pTZ/Sal (Fig. 1).

Preparation of pTZ/Sal ssDNA and site-directed mutagenesis of dengue virus DNA sequences. The vector pTZ18U contains the f1 origin of replication in opposite transcriptional orientation to the T7 RNA polymerase promoter and to the ampicillin resistance gene. To produce pTZ/Sal single-stranded DNA (ssDNA), we followed a standard protocol. To produce mutations in dengue virus ssDNA, we employed the site-directed method of Zoller and Smith (35). Klenow fragment was obtained from Promega Biotec, Madison, Wis., and T4 DNA ligase was obtained from New England BioLabs, Inc., Beverly, Mass. Oligonucleotide primers were prepared on an Applied Biosystems 380A DNA Synthesizer. The presence of a desired mutation was confirmed by primer-directed DNA sequencing of the pertinent region in dengue virus DNA (21).

Synthesis of RNA transcripts from recombinant vector DNA. pTZ/Sal DNA and cloned mutant recombinant DNAs were linearized by digestion with the restriction endonuclease *Hin*dIII (Fig. 1). Synthesis of RNA on 2 to 5 μg of linearized DNA was carried out by a standard procedure using 20 to 40 U of T7 RNA polymerase (Promega Biotec).

Cell-free translation of RNA transcripts and polyacrylamide gel analysis of products. Cell-free translation was conducted in a 25- or 50-μl final volume using a supplemented rabbit reticulocyte lysate pretreated with micrococcal nuclease (Promega Biotec). A typical 25-μl reaction mixture contained 17.5 μl of rabbit reticulocyte lysate, 0.5 μl of a 1 mM amino acid mixture minus methionine, 0.5 μl of RNase inhibitor (RNasin; Promega Biotec), 2.5 μl of [^{35}S]methionine (15 mCi/ml, >1,000 Ci/mmol; Amersham Corp., Arlington Heights, Ill.), and 0.2 to 0.5 μg of RNA in water. Protein synthesis was conducted in the presence or absence of added canine pancreatic microsomes (1.0 to 1.5 μl replacing water in a 25-μl reaction mixture; Promega Biotec). To prepare proteins for sequencing, 1 mM amino acids minus methionine and leucine was added to the reaction mixture in place of amino acids minus methionine, and the reaction was carried out in the presence of 25 μCi of [^{3}H]leucine (1 mCi/ml, 150 Ci/mmol; Amersham) and [^{35}S]methionine. Synthesized proteins were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Gels contained 0.2% SDS and 15 to 17% acrylamide. The 17% gels were electrophoresed at 140 constant volts for 16 h at room temperature, dried, and exposed to XAR-2 film (Eastman Kodak Co., Rochester, N.Y.).

Protease and endo F digestion of proteins. To digest proteins with thermolysin (Boehringer Mannheim Biochemicals, Indianapolis, Ind.), 5 μl of a cell-free synthetic reaction mixture was added to 5 μl of TC buffer (10 mm Tris hydrochloride [pH 7.5], 10 mM $CaCl_2$), 1 μl of thermolysin (1-mg/ml solution in TC buffer) was added, and the reaction mixture was incubated for 60 min at 4°C. Thermolysin activity was terminated by the addition of a 10-fold excess of EDTA.

Endoglycosidase F (endo-β-*N*-acetylglucosaminidase F [endo F]) was obtained from Boehringer Mannheim. Two to five microliters of a cell-free synthetic reaction mixture or >5,000 cpm of an eluted [^{35}S]methionine-labeled protein was incubated in 25 μl of endo F buffer (Boehringer Mannheim) for 1 to 3 h at 37°C in the presence of 0.1 to 0.2 U of enzyme.

Immune precipitation of proteins. Antibody to peptide 30 (a gift of R. Houghten), a 15-amino-acid peptide representing the predicted N terminus of the dengue virus type 4, strain 814669 E protein (MRCVGVGNRDFVEGV; residues 280 to 294 [33]), was prepared in rabbits after conjugation to keyhole limpet hemocyanin. This antiserum had an antipeptide titer of 1:1,280 by enzyme-linked immunosorbent assay. Proteins synthesized in vitro were incubated for 16 h at 4°C in the presence of antipeptide antibody at dilutions of 1:10 to

1:50. The antigen-antibody complexes were then collected on staphylococcal protein A-Sepharose CL4B beads (Pharmacia). The beads were washed in RIPA buffer (0.1 M Tris [pH 7.5], 0.15 M NaCl, 1% sodium deoxycholate, 1% Triton X-100, 0.1% SDS) and boiled in Laemmli sample buffer (0.1 M Tris [pH 6.8], 10% glycerol, 1% SDS, 0.2 M 2-mercaptoethanol, 0.001% bromphenol blue) prior to SDS-PAGE.

Amino-terminal amino acid sequencing of proteins. Proteins doubly labeled ([^{3}H]leucine and [^{35}S]methionine) as described above were separated by SDS-PAGE on a gel that had been preelectrophoresed in the presence of running buffer plus 1 mM mercaptoacetic acid. A 2% agarose stacking gel was used instead of acrylamide, and 0.1 mM mercaptoacetic acid was added to the running buffer. Proteins were eluted from polyacrylamide and sequenced in a Beckman Model 890M Protein Sequencer.

RESULTS

Cell-free translation of wild-type RNA transcripts. *Sal* DNA in the pTZ/Sal construct includes the 13 3′-terminal nucleotides of the dengue virus 5′ noncoding region and sequences encoding the capsid (amino acids 1 to 113), prM (amino acids 114 to 279), and the first 23 amino acids of E (amino acids 280 to 302) in the dengue virus serotype 4 precursor polyprotein (33). RNA transcripts prepared from *Hin*dIII-linearized pTZ/Sal DNA were expected to include 21 nucleotides of the pTZ polylinker 5′ to dengue virus-specific sequences as well as 5′-terminal GGG appended to all transcripts by T7 RNA polymerase (8). In addition, transcripts included polylinker sequences 3′ to the dengue virus insert which encode a common C terminus (Gln-Ala-Cys-Lys) in cell-free translation products (Fig. 1).

The products of cell-free translation of wild-type (wt) RNA transcripts are shown in Fig. 2A. [^{35}S]methionine-labeled proteins were directly applied to the gel for SDS-PAGE under reducing conditions. In the absence of added microsomal membranes (Fig. 2A, lane 0), a single major product was evident, with a molecular size of approximately 33 kilodaltons (kDa). This is in agreement with the predicted size of the uncleaved translation product of full-length pTZ/Sal RNA (306 amino acids; estimated molecular mass, 33,660 daltons). Minor lower- and higher-molecular-mass species were also detected. Higher-molecular-mass species are likely to be the product of RNA transcribed from contaminating uncleaved circular pTZ/Sal DNA. Lower-molecular-mass proteins could have resulted from internal initiation of translation or from internal cleavage of the 33-kDa polypeptide. Evidence favored internal initiation; separate experiments failed to demonstrate a precursor-product relationship between the 33-kDa and lower-molecular-mass proteins, suggesting that cleavage of the full-length polypeptide did not occur in the absence of microsomes (data not shown).

When RNA transcripts were translated in the presence of canine pancreatic microsomal membranes, seven proteins were detected. Four high-molecular-mass products (A, B, and C in Fig. 2A, lanes 0.5 to 2., and C′ in Fig. 3A, left lane) ranged in size from 32 to 38 kDa. Protein C appeared to be identical in size to the unglycosylated 33-kDa full-length product. Three low-molecular-mass proteins varied in size from about 24 to 28 kDa (a, b, and c in Fig. 2A, lanes 0.5, 1., and 2.). Increasing portions of membranes added to the reaction mixture were associated with both enhanced production of proteins a, b, and c relative to the precursor and diminished total synthesis.

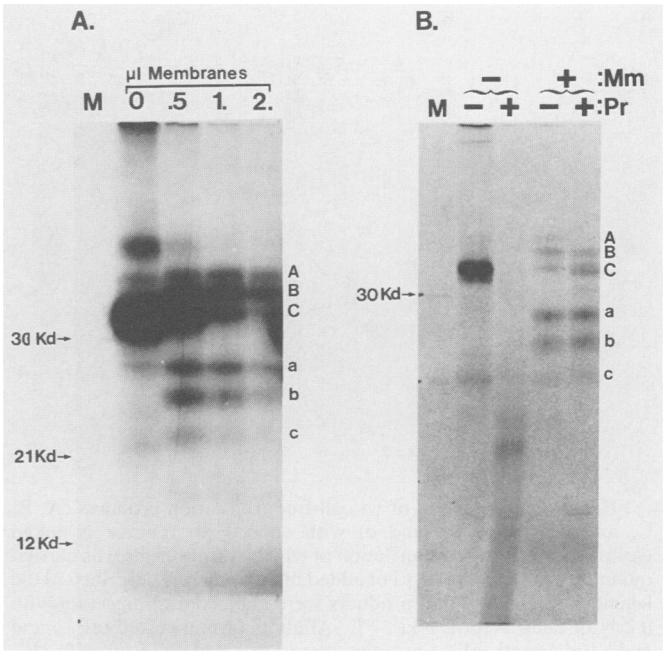

FIG. 2. (A) Cell-free translation of *Sal* RNA in the absence (lane 0) or presence of increasing portions (in microliters) of microsomal membranes (lanes 0.5 to 2.). One-half microgram of RNA transcribed in vitro was added to 17.5 μl of micrococcal nuclease-treated rabbit reticulocyte lysate plus 0.5 μl of RNasin, 0.5 μl of 1 mM amino acids minus methionine, 2.5 μl of [^{35}S]methionine, and microsomes as indicated, to a final volume of 25 μl. The cell-free translation reaction mixture was incubated for 60 min at 30°C. Three-microliter samples of each reaction mixture were added to Laemmli sample buffer, heated to 90°C for 3 min, and directly applied to a 17% gel for SDS-PAGE. ^{14}C-labeled low-molecular-mass marker proteins were co-electrophoresed (lane M). The gel was dried and exposed to Kodak XAR-2 film. Kd, Kilodaltons. (B) Protease sensitivity of in vitro translation products. Cell-free translation was carried out as described above in the absence (−) or presence (+) of microsomes (Mm). After incubation, aliquots of each reaction mixture were subjected to digestion with thermolysin for 60 min at 4°C in TC buffer (Pr+). Protease-digested aliquots were coelectrophoresed with aliquots of undigested product (Pr−). Electrophoresis was done as described in the legend to panel A. M, ^{14}C-labeled marker proteins.

Translocation of cell-free translation products. Resistance to protease digestion of cell-free translation products under conditions that maintained the structural integrity of the endoplasmic reticulum demonstrated that they had been translocated (Fig. 2B). As a control, the 33-kDa full-length polypeptide synthesized in the absence of membranes was completely protease sensitive. Under the same conditions, membrane-dependent products B, C, a, b, and c were protected from protease digestion. Protein C′ was not detected in this experiment. Protein A appeared partially protease sensitive. Protease digestion removed protein A and introduced a new high-molecular-mass product that migrated slightly slower than protein C. This result could be explained by a transmembrane orientation of protein A which renders an extralumenal piece susceptible to protease.

Digestion of glycoproteins with endo F. To detect N-linked glycosylation of translocated proteins, the mixture of A, B, C, and C′ with a, b, and c was digested with the enzyme endo F (Fig. 3A). The products of digestion were four polypeptides designated 1, 2, 3, and 4 in descending order of

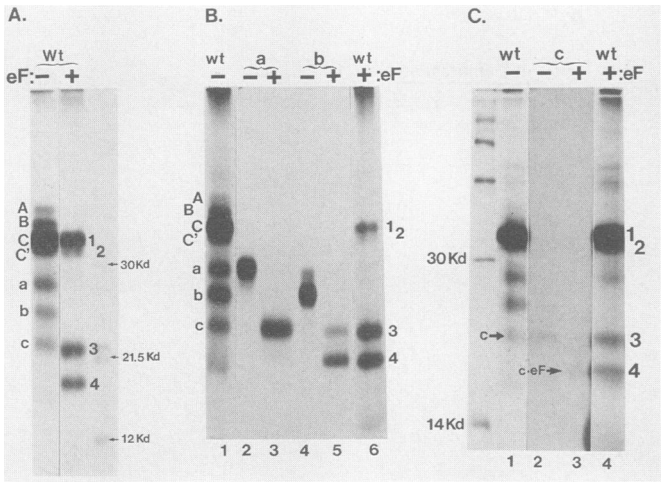

FIG. 3. (A) Digestion of wt cell-free translation products (A, B, C, and C′ plus a, b, and c) with endo F to remove N-linked carbohydrate. In vitro translation of wt RNA transcripts was carried out in the presence of 1.5 μl of added microsomes as described in the legend to Fig. 2A. Total products were subjected to digestion with 0.2 U of endo F for 2 h at 37°C. Aliquots of undigested (eF−) and endo F-digested (eF+) proteins were coelectrophoresed with ^{14}C-labeled marker proteins as in Fig. 2. (B) Identification of cell-free translation products a and b in relation to endo F digestion. wt cell-free translation products prepared as described in the legend to Fig. 2A were separated on a preparative gel. Bands representing products a and b were located by direct exposure to Kodak XAR-2 film, excised, and electroeluted. Products a and b were then separately digested by endo F. The electrophoretic mobilities of proteins a and b (lanes 2 and 4) and their endo F products (lanes 3 and 5) were compared with total proteins (lane 1) and endo F-digested total proteins (lane 6) during SDS-PAGE. (C) Identification of cell-free translation product c in relation to endo F products. [^{35}S]methionine-labeled protein c was prepared, separated from other cell-free translation products, digested by endo F, and compared with total wt and endo F-digested wt proteins as described above. Lane 1, wt products (c is indicated); lane 2, isolated c; lane 3, endo F-digested c; lane 4, total wt proteins digested with endo F. The identity of endo F polypeptide 4 with endo F-digested product c (c · eF) is indicated. ^{14}C-labeled marker proteins were coelectrophoresed (left lane). kd, Kilodaltons.

molecular size (33, 31, 22, and 19 kDa, respectively). These were consistent with the expected sizes of unglycosylated polypeptides representing the full-length precursor, capsid + prM, prM + E, and prM. A polypeptide consistent in size with that of the cleaved capsid (12 to 14 kDa) was not detected. Translation products were next separately eluted from a preparative gel and digested with endo F. Proteins A, B, C, and C′ collectively yielded polypeptides 1 and 2. Proteins A and B were reduced in size by endo F digestion, showing that they were glycosylated. Proteins C and C′ appeared identical to polypeptides 1 and 2 (data not shown). Glycosylation of proteins a, b, and c was also demonstrated (Fig. 3B and C). Glycoprotein a yielded polypeptide 3. Glycoproteins b and c both yielded polypeptide 4. Glycoproteins a and b were each approximately 6 kDa larger than polypeptides 3 and 4, respectively. Glycoprotein c was 3 kDa larger than its endo F digestion product, polypeptide 4. It appeared that endo F removed two carbohydrate residues ($GlcNAc_2$-Man_9-Glc_3, where GlcNAc is *N*-acetylglucosamine, Man is mannose, and Glc is glucose; M_r, 3 kDa) each from proteins a and b and one from c. In the truncated precursor, potential glycosylation sites lie entirely within the prM sequence (discussed elsewhere).

Sequencing and immune precipitation of endo F digestion products. The N and C termini of endo F digestion products 1 through 4 were subsequently identified. Initially, the N-terminal 30 residues of polypeptides that had been labeled with both [^{35}S]methionine and [^{3}H]leucine were sequenced. By alignment of the labeled residues, the N-terminal sequences of polypeptides 1 and 2 were demonstrated to be identical to that predicted for the dengue virus capsid protein (data not shown). Sequence data for polypeptide 3 is shown in Fig. 4A. Within the first 30 residues, methionine was detected in position 12 and leucine was detected in positions 3, 11, 23, and 24. The results of sequencing the first 30 residues of polypeptide 4 were identical. The locations of methionine and leucine residues within the N termini of polypeptides 3 and 4 were in agreement with those predicted for the N-terminal sequence of the dengue virus prM (2, 4, 5, 18, 33), showing that specific cleavage at the capsid-prM site occurred when microsomes were present during cell-free translation.

To detect cleavage of the first 23 amino acids of E from the C terminus of the truncated precursor, we prepared rabbit antibody to a peptide representing the 15 N-terminal residues of E and demonstrated that antipeptide antibody precipitated E from a lysate of dengue virus-infected cells. The same antibody preparation specifically immune precipitated polypeptides 1 and 3 (Fig. 4B, lane 3), indicating that each retained the C-terminal E sequences. Precipitation by preimmune serum is shown as a negative control (lane 4). In contrast, polypeptides 2 and 4 were not immune precipitated by the antipeptide antibody and appear to represent species cleaved at the prM-E site. The endo F polypeptides were therefore identifiable by apparent molecular size, N-terminal sequence, and C-terminus-specific immune precipitation (Table 1): polypeptide 1 = capsid + prM + E (the full-length precursor); polypeptide 2 = capsid + prM; polypeptide 3 = prM + E; and polypeptide 4 = prM. Consequently, proteins a, b, and c were verified as specific glycosylated cleavage products of the 33-kDa precursor. Protein a was identified as prM + E bearing two carbohydrate residues. Proteins b and c were identified as alternate glycosylated forms of prM. Cleavage was membrane dependent at both the capsid-prM and prM-E sites, and cleavage products were translocated. These results confirm and extend data which suggested that the signal peptidase is responsible for cotranslational cleavage of the flavivirus structural proteins (25); however, they do not clarify the significance of the conserved potential proteolytic site in the capsid sequence or localize signal function in the truncated precursor. Experiments that follow address these questions.

Cell-free translation of transcripts bearing mutations in the capsid sequence. Two in-frame deletion mutations were introduced into the capsid nucleotide sequence in pTZ/Sal DNA by site-directed mutagenesis to test the possibility that cleavage of the capsid at the conserved potential proteolytic site (Arg-Lys-Arg; residues 97 to 99) is required in association with signal recognition. The mutant *dl*C polypeptide lacked an internal 74 amino acids of the capsid sequence (residues 13 to 86), which should have abrogated any autoproteolytic activity of the capsid or any conformational requirement for normal processing. Mutant *dl*C-pM lacked the trypsinlike site itself (residues 97 to 99; Fig. 5). In the absence of microsomes, the predominant product of *dl*C transcripts was a protease-sensitive 25-kDa polypeptide, consistent in size with that expected for the full-length unglycosylated product of the nucleotide sequence (Fig. 6A). Also noted as a minor species was a 22-kDa protein, the

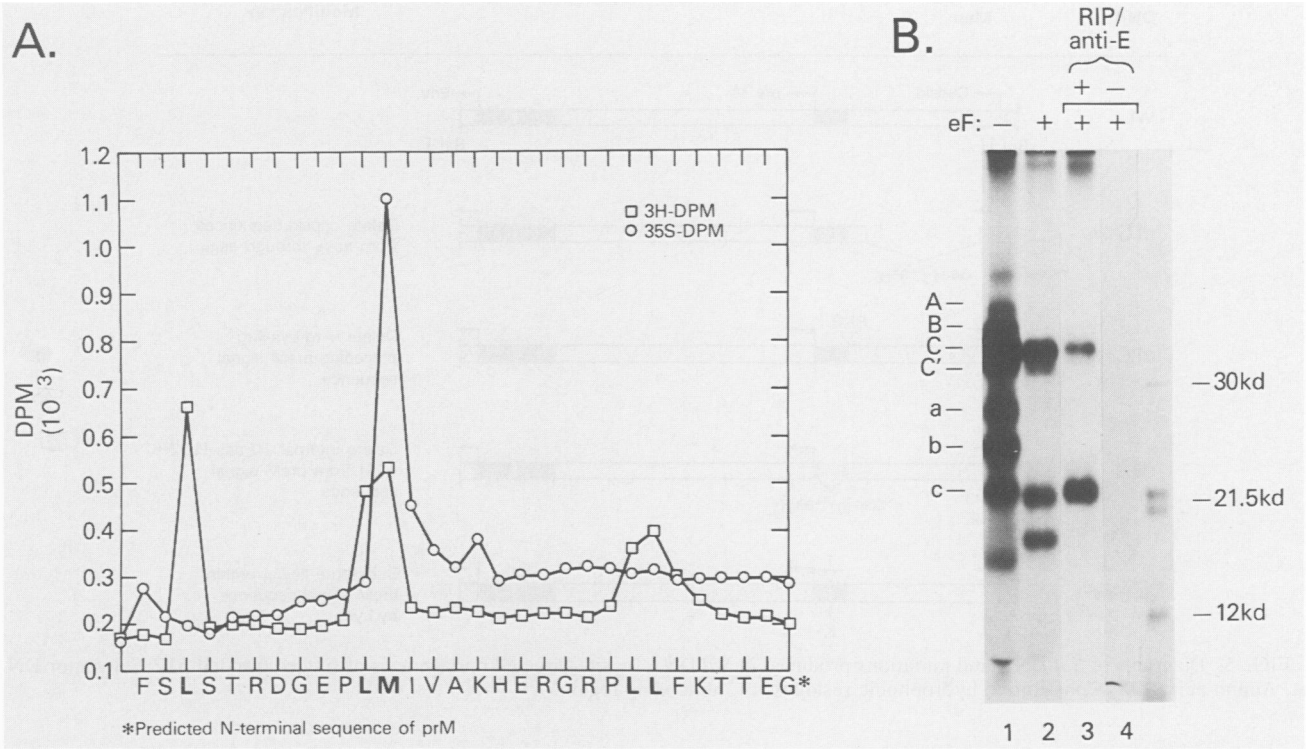

FIG. 4. (A) Sequencing of polypeptide 3. Cell-free translation was carried out in the presence of membranes and [^{3}H]leucine and [^{35}S]methionine labels. The reaction products were digested with endo F and separated on a modified 17% SDS-polyacrylamide gel. Polypeptide 3 was located by exposure of the gel to Kodak XAR-2 film and excised. The protein was electroeluted, and a sample containing 2.7×10^4 dpm of ^{3}H and 4.9×10^4 dpm of ^{35}S was fractionated on a Beckman Model 890M Protein Sequencer. Thirty residues were sequenced from the N terminus (radiolabel in fractions 1 to 30 is shown, left to right). The results are compared with the predicted sequence of the N terminus of the dengue virus type 4 strain 814669 prM (33). The first fraction is aligned with the N-terminal F (single-letter code), and labeled M and L residues are indicated in boldface. (B) Immune precipitation of endo F-digested cell-free translation products by antibody specific for C-terminal E sequences. [^{35}S]methionine-labeled cell-free products (lane 1) were digested with endo F (lane 2). A sample was incubated for 16 h at 4°C with rabbit antibody to a 15-amino-acid peptide representing the predicted N terminus of E (lane 3) or with preimmune serum (lane 4). The immune precipitates were collected by using staphylococcal protein A-Sepharose CL4B beads, eluted from the beads, and coelectrophoresed with undigested and digested total proteins. eF−, Not digested with endo F; eF+, digested with endo F; anti-E+, radioimmunoprecipitation (RIP) with rabbit antipeptide antibody; anti-E−, RIP with preimmune rabbit serum. kd, Kilodaltons.

expected size of a product initiating internally at Met-112 of the wt capsid sequence. The Met-112 codon lies within a consensus nucleotide sequence favoring initiation (12). In the presence of microsomes, translation of mutant *dl*C transcripts produced three translocated proteins which comigrated with the products of wt transcripts (proteins a, b, and c) already identified as glycosylated forms of prM + E and prM. Similarly, the three low-molecular-mass cell-free translation products of mutant *dl*C-pM transcripts in the presence or absence of microsomes were not distinguishable from those of the wt by size or protease sensitivity (Fig. 6B; compare to Fig. 1B). To confirm their identity, processed mutant *dl*C and mutant *dl*C-pM products were subjected to endo F digestion and subsequent immune precipitation by rabbit antibody to E sequences at the C terminus of the precursor (data not shown). This analysis demonstrated that *dl*C and *dl*C-pM products comigrating with wt glycoproteins a, b, and c were identical to them. In conclusion, cleavage of the capsid and glycosylation and translocation of capsidless proteins were not detectably altered by the capsid deletion mutations.

TABLE 1. endo F-digested products of cell-free translation of wt RNA transcripts

Translation product(s)	endo F polypeptide	Molecular size (kDa) by SDS-PAGE	N-terminal sequence	RIP[a] by anti-E	Identity
A, B, C, C′	1	33–34	Capsid	+	C[b] + prM + E
	2	31–32	Capsid	−	C + prM
a	3	22	prM	+	prM + E
b, c	4	19	prM	−	prM

[a] RIP, Radioimmunoprecipitation.
[b] C, Capsid.

Cell-free translation of transcripts bearing mutations in the prM signal. To identify the signal sequence in the truncated precursor, two mutant DNAs were prepared with alterations in the candidate signal (residues 100 to 113) (Fig. 5). Mutant *dl*ss DNA had an in-frame internal deletion of sequences encoding 10 amino acids of the hydrophobic segment (residues 102 to 111). Mutant *sub*ss bears a substitution of the codon for Ile at position 109 by a codon for Lys, which interrupts the hydrophobic sequence. The results of cell-free translation of *dl*ss transcripts compared with wt are shown in Fig. 7A, and the results of protease digestion of *sub*ss products compared with wt are shown in Fig. 7B. The phenotypes of these two mutants were indistinguishable. The signal-minus precursor synthesized in the presence of microsomes was not protected from protease degradation, as

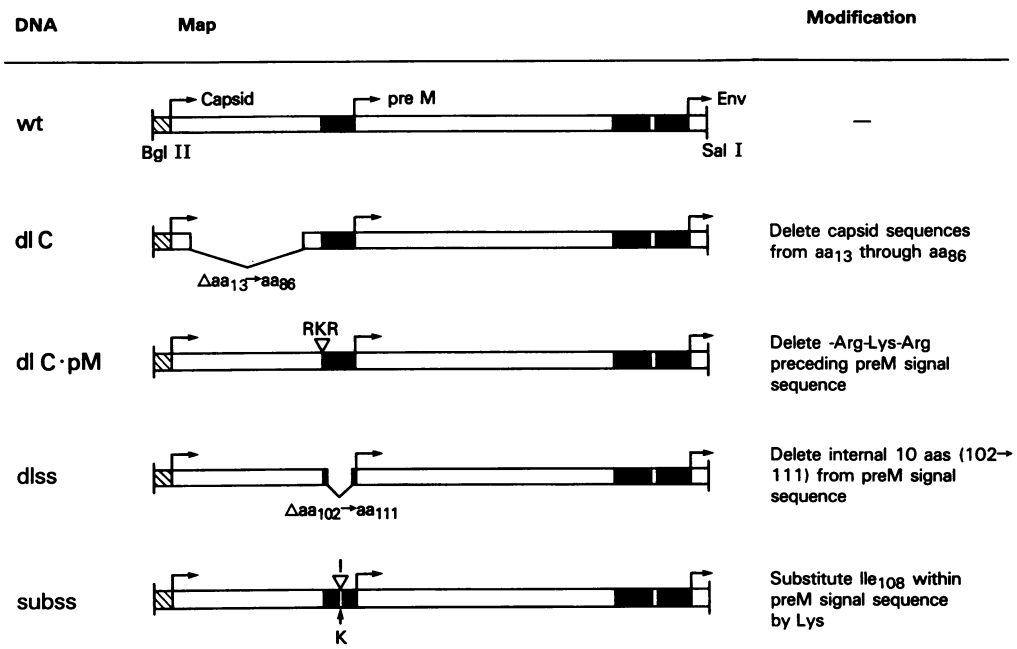

FIG. 5. Diagrams of *Sal* DNA and mutations produced in *Sal* DNA by site-directed mutagenesis of single-stranded pTZ/Sal vector DNA. aa, Amino acid. ■, Consecutive hydrophobic residues; ▨, noncoding region.

shown for *sub*ss. Failure of glycosylation and cleavage is shown for both *dl*ss and *sub*ss. The 33-kDa truncated precursor was unaltered, despite the presence of microsomes in the translation reaction mixture. Interruption or deletion of the hydrophobic segment at once prevented translocation, glycosylation, and cleavage of the precursor at either potential site. These results confirmed a signal function for the hydrophobic segment at the capsid-prM juncture.

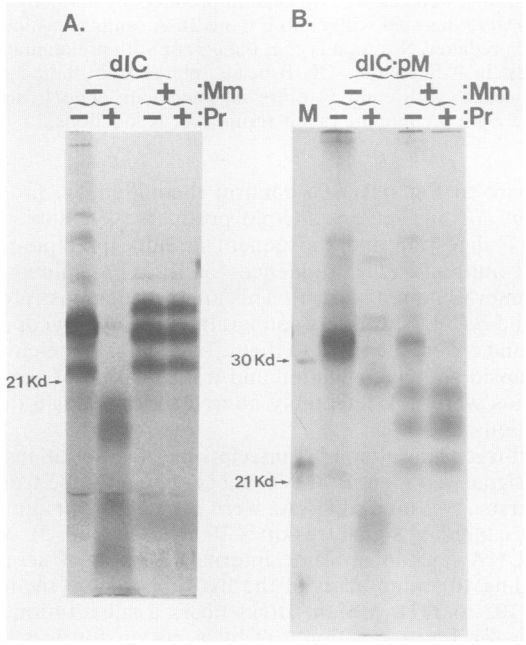

FIG. 6. Cell-free translation of *dl*C (A) and *dl*C-pM (B) mutant RNA transcripts and protease sensitivity of the products. Translation was carried out as described in the legend to Fig. 2A in the presence (+) or absence (−) of 1.5 μl of microsomes (Mm). Thermolysin digestion was carried out as described in the legend to Fig. 2B. Pr−, Undigested products; Pr+, thermolysin-digested products. Products were coelectrophoresed with ^{14}C-labeled standards (lane M), shown in panel B. Kd, Kilodaltons.

DISCUSSION

In vitro translation of RNA transcripts prepared from cloned dengue virus DNA encoding the capsid, prM, and the

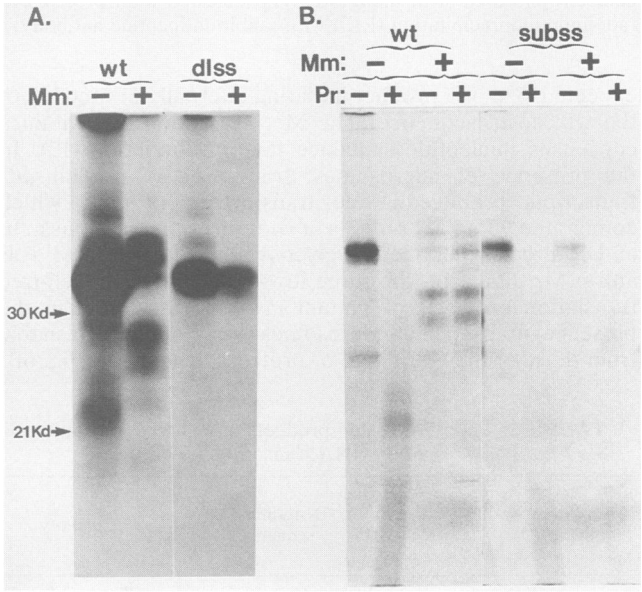

FIG. 7. Cell-free translation of *dl*ss (A) and *sub*ss (B) mutant RNA transcripts and protease sensitivity of the *sub*ss product. Conditions were as described in the legend to Fig. 6. Mm−, Microsomes absent; Mm+, 1.5 μl of microsomes present; Pr−, products not digested with thermolysin; Pr+, products digested with thermolysin. Kd, Kilodaltons.

N-terminal 23 amino acids of E resulted in the synthesis of a precursor polypeptide which was cleaved. Cleavage products representing the capsid + prM, prM + E, and prM were identified, and glycosylated forms of prM + E and prM bearing two carbohydrate residues were detected (proteins a and b). The 33-kDa in vitro precursor contained one consensus site (Asn-X-Ser/Thr) for the addition of N-linked carbohydrate, at Asn-182 in prM. However, prM contains an additional site, Asn-X-Cys (in the conserved sequence Asn-Lys-Cys-Thr; residues 145 to 148), which has been shown to have carbohydrate acceptor function in vitro (1) and in vivo (23). We presume that both Asn residues 145 and 182 are glycosylated in proteins a and b. A third glycosylated cleavage product, prM bearing a single carbohydrate residue, was also detected (protein c). Cleavage of the precursor was wholly dependent on the presence of microsomes during translation. Evidence that in vitro cleavage events are a model of in vivo processing of the dengue virus precursor polyprotein was provided by sequence analysis of the N terminus of prM produced in vitro; its sequence was identical to the predicted N terminus of prM deduced by comparing the dengue virus sequence to those of other flaviviruses for which the N-terminal sequence of prM has been directly determined (2, 4, 33). The approximate accuracy of the cleavage event at the prM-E site was established by immune precipitation of endo F-digested products with an antibody directed against the predicted first 15 amino acids of E (33). The capsid-prM and prM-E cleavage sites in the context of hydrophobic sequences that precede each (5, 18, 19, 33) are highly favored sites for the action of signal peptidase (27), the presumed cleavage enzyme.

To test the hypothesis that signal peptidase is uniquely responsible for cleavage of the capsid and prM and to identify the signal, the phenotype of mutant dengue virus DNA containing either an altered capsid sequence or signal sequence (residues 100 to 113) was determined in cell-free translation experiments. The capsid mutations were generated to rule out the possibility that the capsid is initially cleaved at or upstream from a pair of basic amino acids (residues 98 and 99) which precede the signal for prM, in association with recognition of that signal. The Sindbis virus core protein, which lies N terminal to pE2 and E1 in the Sindbis virus structural polyprotein, initiates processing by a similar cleavage event which is autocatalyzed (9). Mutants with a large internal deletion of capsid sequences (mutant *dl*C) or a deletion of the putative cleavage site (mutant *dl*C-pM) were analyzed. The mutation in the *dl*C polypeptide should have abolished any enzymatic activity of the capsid and removed alternate trypsinlike cleavage sites upstream from residues 98 and 99. However, neither mutation of the dengue virus capsid had a detectable effect on processing of the precursor polypeptide, suggesting that cleavage of the capsid by an activity other than that of signal peptidase is not a prerequisite for processing. As a consequence, it would be predicted that the signal segment is the C terminus of the mature capsid protein. Signallike function of the candidate segment (residues 100 to 113) was established by analysis of mutant polypeptides with a large internal deletion of hydrophobic residues (mutant *dl*ss) or with an interruption in the hydrophobic sequence by an internal substitution of a single hydrophilic amino acid for a hydrophobic amino acid (Ile → Lys; mutant *sub*ss). Both mutations of the putative signal prevented translocation of the polypeptide. Failure of translocation was associated with failure of processing, suggesting that access of intralumenal signal peptidase to the precursor was required for cleavage. Both mutant polypeptides contained the hydrophobic segment at the C terminus of prM (residues 246 to 279) thought to be the signal for E (14, 18, 22, 33). This segment did not mediate translocation of the truncated precursor polypeptide encoded by pTZ/Sal transcripts, perhaps because the sequences downstream (residues 279 to 302) are insufficient in length to drive translocation (17) or because, in fact, translocation of E is dependent on the signal for prM. These findings together suggest that processing of the flavivirus precursor is initiated by cotranslational recognition of the prM signal segment. Translocation affords access of intralumenal signal peptidase to hydrophobic segments preceding the N termini of both prM and E. Signal peptidase effects cleavage at the predicted sites.

The cleaved capsid protein was not detected in these experiments. The capsid should have been labeled at a reduced specific activity with respect to prM, since sequence data predict that it contains only 5 methionines among 113 residues, whereas prM contains 10 methionines among 166 residues. This disparity alone would probably not account for our failure to locate the capsid among labeled translation products. Once cleaved, the capsid may be unstable due to proteolysis or may be poorly resolved under the electrophoretic conditions employed. Chimeric glycoproteins containing the capsid were identified, and these appeared to be translocated (proteins A and B). The finding of uncleaved products containing the capsid and prM may be an artifact of inefficiency of cleavage during translocation in vitro. However, a chimera of the capsid and prM, the 30-kDa glycoprotein NVX, has been detected among flavivirus proteins synthesized during infection (13, 29, 30), suggesting that cleavage of the capsid may also be inefficient in vivo.

ACKNOWLEDGMENTS

I thank Bangti Zhao and Ching-Juh Lai for the generous gift of cloned dengue virus DNA, Suzanne Emerson and Peter L. Collins for many helpful discussions, John Coligan and Michael Raum for protein sequencing, Myron Hill for preparation of oligonucleotides, and Christina Fonseca and Linda Jordan for editorial assistance.

LITERATURE CITED

1. **Bause, E., and G. Legler.** 1981. The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr (Ser) for the N-glycosylation step during glycoprotein biosynthesis. J. Biochem. **195**:639–644.
2. **Bell, J. R., R. M. Kinney, D. W. Trent, E. M. Lenches, L. Dalgarno, and J. H. Strauss.** 1985. Amino-terminal amino acid sequences of structural proteins of three flaviviruses. Virology **143**:224–229.
3. **Boulton, R. W., and E. G. Westaway.** 1972. Comparison of togaviruses: Sindbis virus (group A) and Kunjin virus (group B). Virology **49**:283–289.
4. **Castle, E. T., U. Leidner, G. Wengler, and G. Wengler.** 1985. Sequence analysis of the viral core protein and the membrane-associated proteins V1 and NV2 of the flavivirus West Nile virus and the genome sequence for these proteins. Virology **145**:227–236.
5. **Coia, G., M. D. Parker, G. Speight, M. E. Byrne, and E. G. Westaway.** 1988. Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene order and characteristics of the virus-specified proteins. J. Gen. Virol. **69**:1–21.
6. **Dalgarno, L., D. W. Trent, J. H. Strauss, and C. M. Rice.** 1986. Partial nucleotide sequence of Murray Valley encephalitis virus: comparison of the encoded polypeptides with yellow fever virus structural and nonstructural proteins. J. Mol. Biol. **187**:309–323.
7. **Deubel, V., R. M. Kinney, and D. W. Trent.** 1986. Nucleotide sequence and deduced amino acid sequence of the structural proteins of dengue type 2 virus, Jamaica genotype. Virology **155**:365–377.

8. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. **166**:477–535.
9. **Hahn, C. S., E. G. Strauss, and J. H. Strauss.** 1985. Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid protein autoprotease. Proc. Natl. Acad. Sci. USA **82**:4648–4652.
10. **Hahn, Y. S., R. Galler, T. Hunkapiller, J. M. Dalrymple, J. H. Strauss, and E. G. Strauss.** 1988. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. Virology **162**:167–180.
11. **Hase, T., P. L. Summers, K. H. Eckels, and W. B. Baze.** 1987. An electron and immunoelectron microscopic study of dengue 2 virus infection of cultured mosquito cells: maturation events. Arch. Virol. **92**:273–291.
12. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**:283–292.
13. **Lyapustin, V. N., Y. V. Svitkin, T. Y. Ugarova, V. A. Lashkevich, and V. A. Agol.** 1986. A tentative model of formation of structural proteins of tick-borne encephalitis virus (flavivirus). FEBS Lett. **200**:314–316.
14. **Mackow, E., Y. Makino, B. Zhao, Y.-M. Zhang, L. Markoff, A. Buckler-White, M. Guiler, R. Chanock, and C.-J. Lai.** 1987. The nucleotide sequence of dengue type 4 virus: analysis of genes coding for nonstructural proteins. Virology **159**:217–228.
15. **Mason, P. W., P. C. McAda, T. L. Mason, and M. J. Fournier.** 1987. Sequence of the dengue-1 virus genome in the region encoding the three structural proteins and the major nonstructural protein NS1. Virology **161**:262–267.
16. **McAda, P. C., P. W. Mason, C. S. Schmaljohn, J. M. Dalrymple, T. L. Mason, and M. J. Fournier.** 1987. Partial nucleotide sequence of the Japanese encephalitis virus genome. Virology **158**:348–360.
17. **Rappoport, T. A., R. Heinrich, P. Walter, and T. Schulmeister.** 1987. Mathematical modeling of the effects of the signal recognition particle on translation and translocation of proteins across the endoplasmic reticulum membrane. J. Mol. Biol. **195**:621–636.
18. **Rice, C. M., E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss.** 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science **229**:726–733.
19. **Rice, C. M., E. G. Strauss, and J. H. Strauss.** 1986. Structure of the flavivirus genome, p. 279–326. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Publishing Corp., New York.
20. **Russell, P. K., and A. Nisalak.** 1967. Dengue virus identification by the plaque reduction neutralization test. J. Immunol. **99**: 291–296.
21. **Sanger, R., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Speight, G., G. Coia, M. D. Parker, and E. G. Westaway.** 1988. Gene mapping and positive identification of the non-structural proteins NS2A, NS2B, NS3, NS4B and NS5 of the flavivirus Kunjin and their cleavage sites. J. Gen. Virol. **69**:23–34.
23. **Stenflo, J., and P. Fernlund.** 1982. Amino acid sequence of the heavy chain of bovine protein C. J. Biol. Chem. **257**:12180–12190.
24. **Strauss, E. G., and J. H. Strauss.** 1985. Assembly of enveloped animal viruses, p. 205–234. *In* S. Casjens (ed.), Virus structure and assembly. Jones and Barlett, Boston.
25. **Svitkin, Y. V., T. Y. Ugawara, T. V. Chemovskaya, V. N. Lyapustin, V. N. Lashkevich, and V. L. Agol.** 1984. Translation of tick-borne encephalitis virus (Flavivirus) genome *in vitro*: synthesis of two structural polypeptides. Virology **110**:26–34.
26. **Trent, D. W., R. M. Kinney, B. J. B. Johnson, A. V. Vorndam, J. A. Grant, V. Deubel, C. M. Rice, and C. Hahn.** 1987. Partial nucleotide sequence of St. Louis encephalitis virus RNA: structural proteins, NS1, NS2a, and NS2b. Virology **156**:293–304.
27. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**:4683–4690.
28. **Wengler, G., G. Wengler, and H. J. Gross.** 1978. Studies on virus-specific nucleic acids synthesized in vertebrate and mosquito cells infected with flaviviruses. Virology **89**:423–427.
29. **Westaway, E. G.** 1973. Proteins specified by group B togaviruses in mammalian cells during productive infections. Virology **51**: 454–465.
30. **Westaway, E. G.** 1987. Flavivirus replication strategy. Adv. Virus. Res. **33**:45–90.
31. **Winkler, G., F. X. Heinz, and C. Kunz.** 1987. Studies on glycosylation of flavivirus E proteins and the role of carbohydrate in antigenic structure. Virology **159**:237–243.
32. **Wright, P. J., H. M. Warr, and E. G. Westaway.** 1981. Synthesis of glycoproteins in cells infected by the flavivirus Kunjin. Virology **109**:418–427.
33. **Zhao, B., E. Mackow, A. Buckler-White, L. Markoff, R. M. Chanock, C.-J. Lai, and Y. Makino.** 1986. Cloning full-length dengue type 4 viral DNA sequences: analysis of genes coding for structural proteins. Virology **155**:77–88.
34. **Zhao, B., G. Prince, R. Horswood, K. Eckels, P. Summers, R. Chanock, and C.-J. Lai.** 1987. Expression of dengue virus structural proteins and nonstructural protein NS1 by a recombinant vaccinia virus. J. Virol. **61**:4019–4022.
35. **Zoller, M. J., and M. Smith.** 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13-derived vectors. Methods Enzymol. **100**:468–500.