JOURNAL OF VIROLOGY, Dec. 1998, p. 10093–10099 Vol. 72, No. 12
0022-538X/98/$04.00+0

Identification and Characterization of the *Escherichia coli*-Expressed RNA-Dependent RNA Polymerase of Bamboo Mosaic Virus

YI-IJA LI,[1] YUEH-MEI CHENG,[2] YIH-LEH HUANG,[1] CHING-HSIU TSAI,[3] YAU-HEIU HSU,[3] AND MENGHSIAO MENG[1]*

Graduate Institute of Agricultural Biotechnology,[1] Graduate Institute of Biochemistry,[2] and Agricultural Biotechnology Laboratories,[3] National Chung Hsing University, Taichung, Taiwan 40227, Republic of China

Received 11 May 1998/Accepted 25 August 1998

Bamboo mosaic virus (BaMV), a member of the potexvirus group, infects primarily members of the *Bambusoideae*. The open reading frame 1 (ORF1) of BaMV encodes a 155-kDa polypeptide that was postulated to be involved in the replication and the formation of cap structure at the 5′ end of the viral genome. To characterize the activities associated with the 155-kDa viral protein, it was expressed in *Escherichia coli* BL21(DE3) cells with thioredoxin, hexahistidine, and S · Tag fused consecutively at its amino terminus, and the fusion protein was purified by metal affinity chromatography. Several RNA fragments, prepared by in vitro transcription, were tested as substrates for the RNA-dependent RNA polymerase (RdRp) activity. Among them, the expressed fusion enzyme was able to generate a ^{32}P-labeled RNA product when 3′-end RNA fragments of the positive strand or negative strand of BaMV were included in the assay mixture. Dot hybridization assay revealed that the reaction products are complementary to their RNA substrates. Taken together, the evidence suggests that the 155-kDa protein encoded by ORF1 of BaMV has an RdRp activity and should be involved in the replication of BaMV. Mutational analyses demonstrate the importance of the GDD motif in the polymerase activity, and deletion studies suggest that the polymerase activity resides in the carboxyl terminus of the 155-kDa viral protein.

Bamboo mosaic virus (BaMV) infects members of *Bambusoideae* throughout Taiwan and adversely affects the economic production of bamboo in Taiwan. The viral genome is a positive-strand RNA molecule with a 5′-^{7}mGpppG cap structure and a 3′ poly(A) tail. As shown in Fig. 1, the 6.4-kb genome [excluding the 3′ poly(A) tail] consists of a 94-nucleotide (nt) untranslated region at the 5′ end, open reading frame 1 (ORF1; 4,098 nt), a triple gene block (ORF2 to ORF4), a coat protein coding region, and a 142-nt 3′-untranslated region (22). ORF1, coding for 1,365 amino acids, was able to direct the synthesis of a 155-kDa protein in an in vitro rabbit reticulocyte lysate system (21). Due to the similarity of its genomic organization to that of other potexviruses, BaMV has been classified as a member of potexvirus group.

An analysis of the amino acid sequence of the 155-kDa protein led to the prediction that this viral protein may contain a methyltransferase, an RNA helicase, and a polymerase. The methyltransferase activity should be located on the N-terminal portion because of the presence of conserved motifs for the putative Sindbis virus-like methyltransferase (30). The helicase should be located in the middle portion that contains several important motifs, including the nucleotide-binding motif GKS (9, 16). The C-terminal portion should be responsible for the polymerase activity, since it contains the GDD motif, a hallmark of polymerase (11, 13–15, 19, 20, 23, 24). Therefore, this 155-kDa protein was postulated as a replicase involved in the formation of a 5′ cap structure and the replication of the viral genome (22).

At present, very little is known about the molecular mechanism of BaMV replication. As with other positive-strand RNA viruses, the replication of BaMV is thought to proceed via the initial synthesis of a complementary negative RNA strand, which in turn serves as a template for the production of progeny positive-strand RNA molecules. As with other potexviruses, the difficulty in obtaining a pure preparation of viral RNA-dependent RNA polymerase (RdRp) in quantity hampers the characterization of the viral enzyme and consequently the studies of BaMV replication. To overcome such obstacles, we set out to develop a heterologous expression system for the production of the 155-kDa viral protein. The inherent activity was identified, and the structural organization of the recombinant viral protein was delineated in the present study as the first step toward the elucidation of the replication mechanism of BaMV in detail.

MATERIALS AND METHODS

Plasmid construction. A 4.1-kb DNA fragment containing all of ORF1 was amplified from a cDNA clone of BaMV (22) by PCR with the *Pfu* polymerase (Stratagene) and the primers 5′-ATGGCACTCGTTTCTAAAGTCTTTGACAGC and 5′-AGTTGTTGAGCTCATAGGTTAC*TTA*ACTAGAGAATAAACC. The first primer has a sequence identical to the first 30 nt of the coding strand of ORF1, while the other primer contains the sequence complementary to the 3′ end of ORF1 with an engineered site for endonuclease *Sac*I (underlined). *TTA* (italicized) in the latter primer represents the complementary sequence of the ORF1 stop codon. PCR was carried out for 28 cycles by a two-temperature process (94°C, 15 s; 63°C, 8 min) followed by a 10-min extension at 72°C in a 50-μl reaction mixture containing a 0.32 μM concentration of each primer, a 0.2 mM concentration of each deoxyribonucleotide, and 2.5 U of *Pfu* polymerase. After digestion with *Sac*I, the amplified DNA fragment was ligated into plasmid pET-32a (Novagen) that had been digested with *Sac*I and *Eco*RV. This expression vector (pEBMN3295) was introduced into *Escherichia coli* BL21(DE3) cells to produce the recombinant viral protein that is believed to consist of thioredoxin, hexahistidine, S · Tag, and the 155-kDa viral protein in order from the N terminus to the C terminus. There is a thrombin cutting site between hexahistidine and S · Tag and an enterokinase cutting sequence between S · Tag and the viral protein.

A PCR-based mutagenesis was used to mutate the GDD motif coding sequence to that encoding GAD. A pair of 5′-phosphorylated divergent primers were annealed to plasmid pEBMN3295 (9.1 kb) at the site of the GDD coding sequence. The sequences of the primers were 5′-**GCC**GATTGCGCCCTCAACTGCGAACCTGAAG and 5′-TCCAGCGTACAGTTGTGCAGTGCCTCT. The boldface sequence is the mutagenic codon coding for alanine. Plasmid

* Corresponding author. Mailing address: Graduate Institute of Agricultural Biotechnology, National Chung Hsing University, 250 Kuo-Kuang Rd., Taichung, Taiwan 40227, Republic of China. Phone: 886-04-2840328. Fax: 886-4-2853527. E-mail: mhmeng@dragon.nchu.edu.tw.

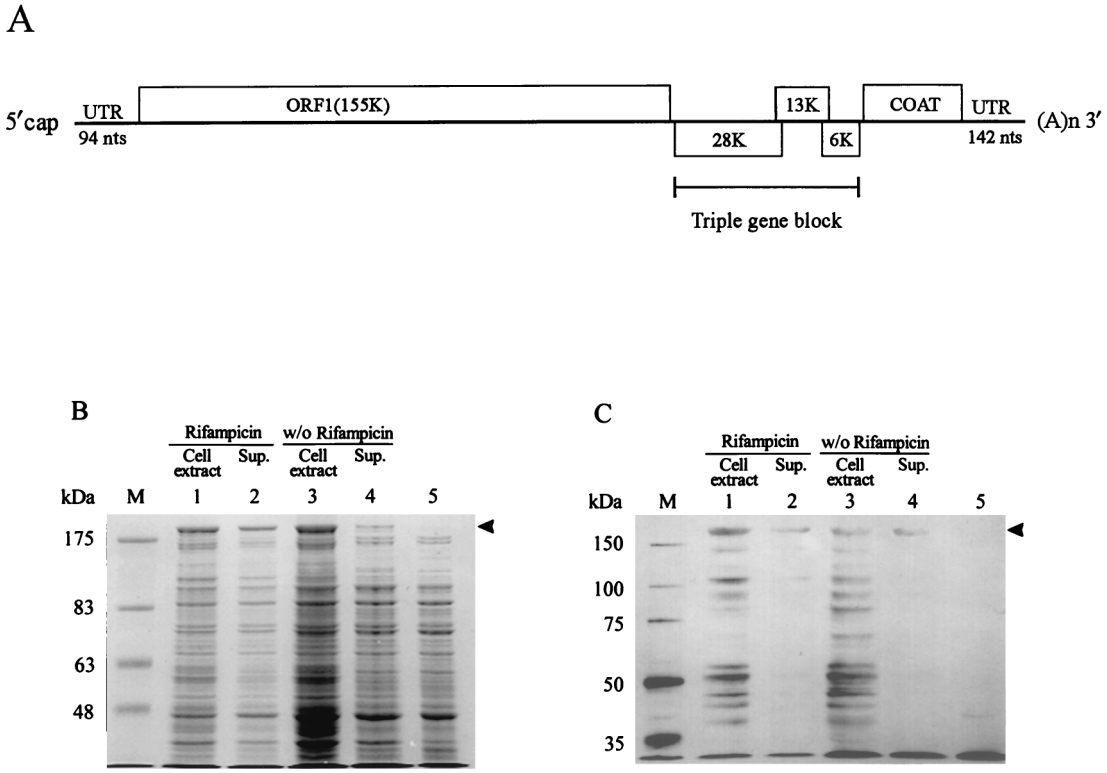

FIG. 1. Expression of the ORF1 product of BaMV in *E. coli* cells. (A) Schematic drawing of the genomic organization of BaMV. The diagram illustrates the locations of the viral genes, each of which is identified by size or function (coat) of the protein product. (B) Analysis of protein expression by SDS-PAGE (8% polyacrylamide). Lanes 1 and 3, cell extracts of *E. coli* cells harboring pEBMN3295. Rifampin was present (lane 1) or absent (lane 3) in the induction process. Lane 2, supernatant after 10 min of centrifugation at 10,000 × *g* of the sample shown in lane 1; lane 4, supernatant of sample shown in lane 3; lane 5, extracts of *E. coli* cells harboring pET-32a. The arrow indicates the full-length fusion protein. M, protein markers with molecular masses of 175, 83, 63, and 48 kDa. (C) Western blot analysis of the recombinant proteins expressed in *E. coli* cells. Alkaline phosphatase-conjugated S protein was used to detect the presence of fusion proteins. The samples from lanes 1 to 5 are the same as those shown in panel B. The marker solution (Perfect Protein Western Markers [Novagen]) contains a mixture of S · Tag recombinant proteins of precise sizes (15 to 150 kDa). The arrow indicates the full-length fusion protein.

pEBMN3295 (5 μg) was subjected to 25 cycles of PCR amplification (94°C, 30 s; 63°C, 16 min) followed by a 20-min extension at 72°C in a 50-μl reaction mixture containing a 0.32 μM concentration of each primer, a 0.2 mM concentration of each deoxyribonucleotide, and 4 U of *Pfu* polymerase. After PCR, the reaction mixture was treated with *Dpn*I to digest the methylated template DNA. The 9.1-kb blunt-ended linear PCR product was then isolated by agarose gel electrophoresis, self-ligated to become a circular plasmid, and transferred into *E. coli* XL1-Blue cells. The mutated plasmid isolated from *E. coli* XL1-Blue cells was subjected to nucleotide sequencing to confirm the desired mutation. To avoid sequencing the whole 4.1-kb fragment of ORF1, a 1.1-kb DNA containing the desired mutation sequence was cut out from the mutated pEBMN3295 with *Eco*RI and *Hin*dIII and then ligated with the larger DNA fragment (8 kb) resulting from the digestion of wild-type pEBMN3295 with the same restriction enzymes. To express the mutated viral protein, *E. coli* BL21(DE3) cells were used as the host cells. The two aspartates in the GDD motif were deleted as described above by using the primers 5′-TGCGCCCTCAACTGCGAACCT and 5′-TCCAGCGTACAGTTGTGCAGTGCC in the PCR-based mutagenesis.

Truncated viral proteins were created by deleting the 5′ portion of ORF1. Primers 5′-ACTAAGAACTGGGAAACAAAACAGGAG and 5′-AGTTGTT GAGCTCATAGGTTAC*TTA*ACTAGAGAATAAACC were used to amplify, from a cDNA clone of BaMV, a deleted DNA fragment (ca. 2.9 kb) that is supposed to code for a truncated protein starting from amino acid 403 and continuing to the C terminus (residue 1365). Primers 5′-AGCGAGGAGCGG AAGTGCC and 5′-AGTTGTTGAGCTCATAGGTTAC*TTA*ACTAGAGAAT AAACC were used to amplify another deletion-containing DNA fragment (ca. 2.6 kb) coding for a truncated protein starting from amino acid 514 and continuing to the C terminus. Another pair of primers, 5′-CTGCCTGACCCGAA GCC and 5′-AGTTGTTGAGCTCATAGGTTAC*TTA*ACTAGAGAATAAA CC, was used to amplify an even shorter DNA fragment (ca. 1.4 kb) coding for amino acids from position 893 to the C terminus. PCR was carried out with *Pfu* polymerase (2.5 U) for 26 cycles (94°C, 30 s; 52°C, 30 s; 72°C, 6 min) in the buffer containing a 0.32 μM concentration of each primer and a 0.2 mM concentration of each deoxyribonucleotide. Each of the amplified DNA fragments was treated with *Sac*I and inserted into plasmid pET-32a that had been treated with *Sac*I and *Eco*RV to become pEBMN321301, pEBMN321634, or pEBMN322771. As with pEBMN3295, these three plasmids were expressed in *E. coli* BL21(DE3) cells to produce the truncated proteins with thioredoxin, hexahistidine, and S · Tag fused at the N termini of the proteins.

Expression and purification of recombinant proteins. To produce the recombinant viral protein, a 2-ml overnight culture of *E. coli* BL21(DE3) cells harboring appropriate plasmids was inoculated into 200 ml of YTGK medium (yeast extracts, 16 g/liter; peptone, 10 g/liter; glycerol, 1%; NaCl, 5 g/liter; and KCl, 0.75 g/liter) and incubated at 37°C. IPTG (isopropyl-β-D-thiogalactopyranoside; final concentration, 0.8 mM) was added to the culture when the cell density reached an optical density at 600 nm of approximately 0.8 to 1 to induce the expression of the lysogenic T7 RNA polymerase that, in turn, transcribed the target gene. Rifampin (final concentration, 100 μg/ml) was added 1 h later to inhibit the *E. coli* RNA polymerase. The cells were harvested by centrifugation after another hour of incubation. The pellets were resuspended in 20 ml of Tris buffer (50 mM, pH 8.0) containing 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and 0.1% Tween 20. This cell suspension at 4°C was sonicated with a sonifier (Vibra Cell; Sonics & Materials, Inc.) with the power (15% amplitude) intermittently on and off in 10-s intervals for 10 min. The cell extract was clarified by centrifugation at 10,000 × *g* for 10 min. To isolate the recombinant viral protein from the crude cell extract, the clarified supernatant was mixed with 1 ml of Talon metal affinity resin (Clontech) and shaken gently at 4°C for 1 h. The mixed resin was then washed 10 times with 2 ml of washing buffer (50 mM Tris buffer, pH 8.0; 150 mM NaCl, 0.1% Triton X-100; 30 mM imidazole) and eluted with 1 ml of elution buffer (50 mM Tris buffer, pH 8.0; 300 mM NaCl; 0.1% Triton X-100; 100 mM imidazole) to obtain the partially purified recombinant protein. The protein eluted from Talon metal affinity resin was then dialyzed in buffer (20 mM Tris, pH 8.0; 100 mM NaCl; 5% glycerol) to remove excess salts.

Protein gel electrophoresis and Western blot analysis. Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and electrotransferred to a polyvinylidene difluoride membrane with a Mini Trans-Blot Cell (Bio-Rad). The filters were incubated with alkaline phosphatase-conjugated S protein directed against S · Tag located on the N-terminal portion of the recombinant protein at a dilution of 1:5,000, and bound S protein

was detected by incubation with substrate solution for alkaline phosphatase (BCIP [5-bromo-4-chloro-3-indolylphosphate toluidinium], 0.25 mg/ml; nitroblue tetrazolium, 0.5 mg/ml; 20 mM Tris, pH 7.5; 150 mM NaCl; 0.05% Tween 20).

Preparation of RNA transcripts. To prepare the 200-nt 5′-terminal fragment of the positive-strand RNA of BaMV, a DNA fragment was amplified first from the cDNA clone of BaMV by PCR with primers 5′-GCTCTAGA*TAATACGACTCACTATA*GAAAACCACTCCAAACGAA and 5′-CTGTGTCTTTGAGCAC. The former primer has the first 19 nt of the positive-strand cDNA of BaMV preceded by a T7 promoter sequence (in italics). The latter primer has the sequence complementary to the positive-strand DNA from nt 185 to 200. PCR was carried out for 35 cycles (94°C, 30 s; 55°C, 30 s; 72°C, 2 min) followed by a 10-min extension at 72°C with *Taq* polymerase. The amplified fragment (5 μg) was then used as the template in a 50-μl in vitro transcription reaction (40 mM Tris, pH 8.0; 4 mM nucleoside triphosphate; 8 mM $MgCl_2$; 2 mM spermidine; 25 mM NaCl; 20 mM dithiothreitol; 10 U of T7 polymerase) to generate the 5′-terminal fragment of the positive-strand RNA. After a 2-h incubation at 37°C, RNase-free DNase I was included in the reaction mixture to digest the template DNA. The RNA transcript was further purified by extracting the RNA from a 7% polyacrylamide gel with elution buffer (0.5 M ammonium acetate, 1 mM EDTA, 10 mM magnesium acetate, and 0.1% SDS), treated once with phenol-chloroform, and precipitated with ethanol. This 200-nt RNA fragment contains the entire 5′-untranslated region and the first 106-nt sequence of ORF1.

Sp6 RNA polymerase was used in the in vitro transcription reaction to generate the 178-nt 3′-terminal fragment of the negative-strand RNA and the 190-nt 3′-terminal fragment of the positive-strand RNA. Otherwise, the conditions for the PCR and in vitro transcription were as described above. The primers 5′-*GATTTAGGTGACACTATAG*AACTTGGGAGTGCGCTTCCTC and 5′-GAAAACCACTCCAAACGAAA were used to generate the 3′-terminal fragment of the negative-strand RNA, and the primers 5′-*GATTTAGGTGACACTATAG*AACGTTGCATGATCGTAAAAC and 5′-GCGGGATCC$(T)_{40}$GGAAAAA were used for the 3′-terminal fragment of the positive-strand RNA. The sequence indicated by italics in the primers represents the SP6 promoter. The 3′-terminal fragment of the positive-strand RNA contains the entire 3′-untranslated region followed by a 40-nt poly(A) stretch and a 9-nt nonviral sequence. The 3′-terminal fragment of the negative-strand RNA is complementary to the 5′ terminus of the positive-strand RNA genome from the first nucleotide to position 178.

To prepare the 334-nt fragment of the satellite C4 RNA of cucumber mosaic virus, the plasmid pCST4-1 (11a) was linearized with *Eco*RI, and a runoff transcription was performed with T7 RNA polymerase.

In the preparation of the 5′ terminus of the negative-strand RNA of BaMV, the plasmid pBL3′(−)sb2.6 (11b), which contains part of the cDNA of BaMV flanked by T7 and T3 promoters, was linearized with *Hin*dIII, and a runoff transcription was performed with T7 RNA polymerase. This 635-nt RNA transcript contains the sequence that is complementary to part of the coat protein coding region and the entire 3′-untranslated region.

The ^{32}P-labeled RNA transcripts were prepared as described above except that 10 μCi of [α-^{32}P]UTP (3,000 Ci/mmol) was included in the transcription reaction mixture (40 mM Tris, pH 8.0; 1 mM ATP; 1 mM GTP; 1 mM CTP; 10 μM UTP; 8 mM $MgCl_2$; 2 mM spermidine; 25 mM NaCl; 20 mM dithiothreitol; and 10 U of RNA polymerase).

In vitro RdRp assay. The reaction was initiated by adding purified enzyme (ca. 5 μg) into a final 50-μl reaction mixture that contained 50 mM Tris (pH 8.0), 2 μg of RNA substrate, 10 mM $MgCl_2$, 10 mM dithiothreitol, 5% glycerol, 20 U of RNase inhibitor, 5 μCi of [α-^{32}P]GTP (3,000 Ci/mmol), 0.25 mM ATP, 0.25 mM CTP, 0.25 mM UTP, and 2.5 μM GTP. Reactions were performed at 26°C for 2 h. An equal volume of stop solution (50 mM Tris, pH 8.0; 10 mM EDTA; 4% SDS) was added to stop the reaction. Five micrograms of yeast tRNA was included in the reaction mixture before the reaction product was extracted twice with phenol-chloroform and precipitated with a 0.5 volume of ammonium acetate (7.5 M) and 2 volumes of ethanol. The dried reaction product was then dissolved in 10 μl of TE buffer and 5 μl of formamide solution (90% formamide, 0.1% bromophenol blue, 0.1% xylene cyanol, 1 mM EDTA) and heated at 95°C for 3 min prior to loading onto 7% polyacrylamide gels containing 7 M urea. The results were visualized by autoradiography.

Dot blot assay. Approximately 100 ng each of the 5′ termini of positive-strand RNAs and the 3′ termini of the negative-strand RNAs was spotted separately onto a strip of nylon paper that had been soaked in 2× SSC buffer (300 mM NaCl, 65 mM sodium acetate; pH 7.0). Immediately thereafter, the RNA transcripts were cross-linked on the paper by using UV light. The RNA-fixed paper was then prehybridized with sodium phosphate buffer (250 mM, pH 7.2) containing 7% SDS at 55°C. One hour later, the ^{32}P-labeled reaction product (400 μl) of the RdRp assay, in which the 3′ terminus of the negative-strand RNA had been used as the RNA substrate, was included in the hybridization buffer, and the reaction was carried out at 55°C for another 16 h. The ^{32}P-labeled 3′ terminus of the negative-strand RNA, which was generated by in vitro transcription, was used as a control to hybridize with another RNA-fixed nylon paper. The ^{32}P-labeled RNA probes were heated at 100°C for 5 min before being loaded into the hybridization mixture. The hybridization solution was then replaced twice with washing buffer I (20 mM sodium phosphate, pH 7.2; 5% SDS) and then twice with washing buffer II (20 mM sodium phosphate, pH 7.2; 1% SDS). Each wash was accompanied by incubation at 55°C for 30 min. The treated nylon paper was air dried, and the results were visualized by autoradiography.

The 3′ terminus of the positive-strand RNA and the 5′ terminus of the negative-strand RNA of BaMV were cross-linked on another set of nylon papers. One RNA-fixed paper was hybridized with the ^{32}P-labeled reaction product (400 μl) of the 3′ terminus of the positive-strand RNA of BaMV, while the other paper was hybridized with the ^{32}P-labeled 3′ terminus of the positive-strand RNA, generated by in vitro transcription, under the same conditions as described above.

RESULTS

Expression and purification of the products of ORF1. Viral RdRp plays a central role in replication of the RNA genome, which is a critical step in the pathogenesis of many RNA viruses. To date, most of the knowledge regarding the biochemical properties of RdRp comes from studies of animal RNA viruses, such as poliovirus (2, 4, 14), hepatitis C virus (3, 23), and rabbit hemorrhagic disease virus (38), and plant viruses, such as brome mosaic virus (17, 31, 33, 35) and tobacco mosaic virus (25). Unlike that of the enzyme from animal viruses, the structural organization of plant viral RdRp is relatively less well characterized. The difficulty in purifying RdRp from the infectious plants and the lack of an efficiently heterologous expression system may account for this delay. To our knowledge, the heterologous expression of plant viral RdRp has been reported only in few cases, such as with the Nlb protein of tobacco vein mottling virus (11). The product of ORF1 of potexviruses may be involved in the formation of a cap structure at the 5′ end and the replication of the RNA genome. To derive biochemical evidence for the function of this viral protein, we established an expression system to produce the ORF1 product of BaMV. The corresponding cDNA fragment was inserted into expression vector pET-32a, and the 155-kDa viral protein was expressed in *E. coli* BL21(DE3) cells as part of a fusion protein with thioredoxin, hexahistidine, and S · Tag fused at its N terminus consecutively. A significant amount of this fusion protein was expressed, as shown in Fig. 1B. The inclusion of rifampin in the induction process lowered the amount of expressed protein but increased the amount of soluble full-length protein. Western blotting with S protein directed against S · Tag confirmed that the protein with the greatest molecular mass should represent the full-length fusion protein (Fig. 1C). S protein also recognized several protein bands with smaller molecular weights than that of the full-length fusion protein. Such smaller protein fragments may represent the immature translational products or may be the proteolytic products of the full-length fusion protein.

Affinity chromatography based on the interaction of S protein and S · Tag was attempted in order to recover the full-length fusion protein. Unfortunately, the recovery rate was low due to the inefficient digestion of the protease cutting sites by thrombin or enterokinase. Talon metal affinity chromatography, which is based on the interaction of hexahistidine and Co^{2+}, was found to be able to recover the full-length fusion protein from crude cell extracts along with several small protein fragments (Fig. 2). Fractionation methods, such as differential centrifugation, ammonium sulfate precipitation, or gel filtration, failed to remove the small fragments from the full-length protein, suggesting that the expressed viral proteins of various sizes might be present together in a heterogeneous manner. The enzyme purified by Talon chromatography was used subsequently for the assay of RdRp activity.

Identification of inherent activity of the recombinant viral protein. The importance of the 3′-untranslated region of positive-strand RNA in the initiation of the replication process has been demonstrated in several viruses, including brome mosaic virus (6), turnip yellow mosaic virus (37), and picornaviruses

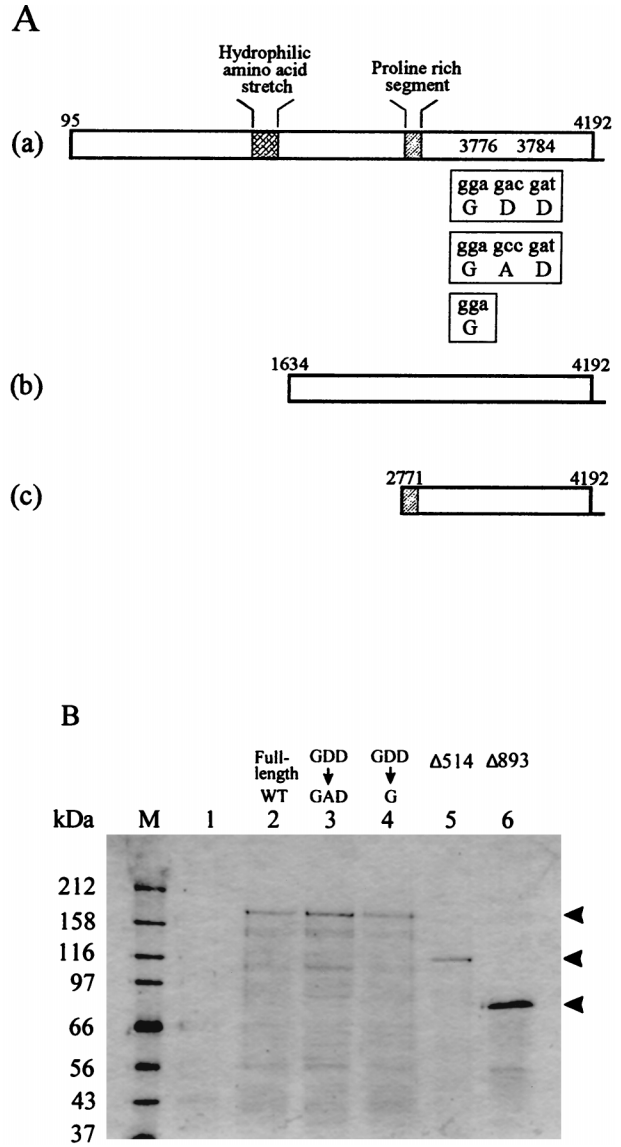

FIG. 2. Purification of wild-type, mutant, and truncated viral proteins. (A) Schematic drawing of the cDNA fragments inserted into pET-32a for the production of wild-type, mutant, and truncated ORF1 protein products. The full-length cDNA of ORF1 is shown diagramatically in part a, with crossed and hatched boxes representing the coding regions of the long hydrophilic amino acid stretch and the proline-rich segment, respectively. The cDNA was mutated around positions 3776 to 3784 to get mutant proteins GDD→GAD and GDD→G (see Materials and Methods). The cDNA sequence is indicated in lowercase letters, and the ultimate effect of the changes on the sequence of the 155-kDa protein is indicated in uppercase letters. The cDNA fragments illustrated in parts b and c encode truncated proteins Δ514 and Δ893, respectively. (B) The proteins produced by *E. coli* cells harboring the appropriate plasmids were purified by Talon metal affinity chromatography. Lane 1, proteins from *E. coli* cells harboring pET-32a; lane 2, wild-type fusion protein from *E. coli* cells harboring pEBMN3295; lane 3, site-directed mutant protein with GDD motif replaced with GAD; lane 4, mutant protein with DD sequence deleted from the GDD motif; lane 5, truncated protein (Δ514) produced by *E. coli* cells harboring pEBMN321634; lane 6, truncated protein (Δ893) produced by *E. coli* cells harboring pEBMN322771.

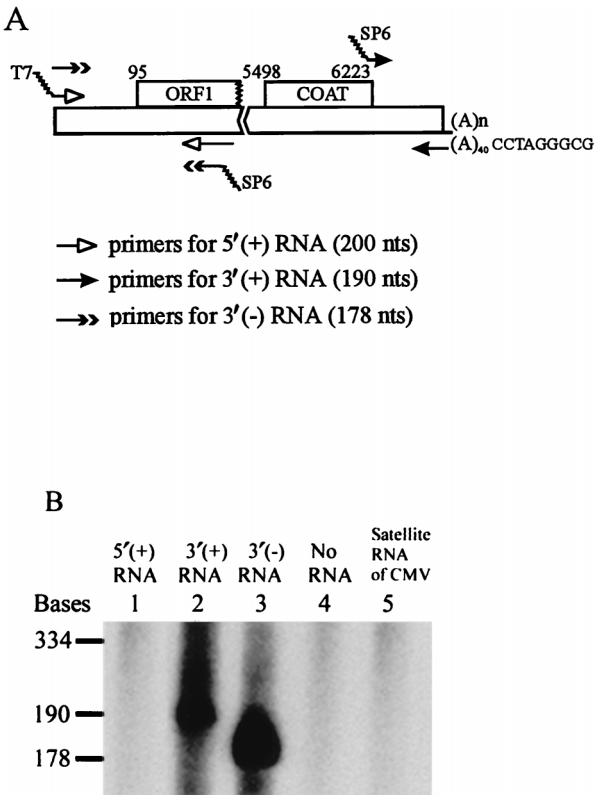

FIG. 3. The RdRp activities of wild-type fusion protein toward various RNA substrates. (A) Schematic drawing of three pairs of primers that were used to generate a promoter (T7 or SP6) containing cDNA fragments by PCR. The amplified cDNA fragments of BaMV were then used to produce the 5′ positive-strand RNA (200 nt), the 3′ positive-strand RNA (190 nt), and the 3′ negative-strand RNA (178 nt) by in vitro transcription (see Materials and Methods). (B) Approximately 5 μg of wild-type enzyme, purified by metal affinity chromatography, was incubated with 2 μg of RNA substrate in an enzyme reaction buffer (see Materials and Methods) at 26°C for 2 h. The reaction products were separated on a 7% polyacrylamide gel containing 7 M urea. Lane 1, activity toward the 5′ positive-strand RNA of BaMV; lane 2, activity toward the 3′ positive-strand RNA of BaMV; lane 3, activity toward the 3′ negative-strand RNA of BaMV; lane 4, RNA not included in the reaction mixture; lane 5, activity toward a 334-nt fragment of satellite C4 RNA of cucumber mosaic virus.

(5, 26). Both the 3′-terminal fragments of positive-strand and negative-strand RNAs of BaMV may assume special structures and be recognized by the product of ORF1 for the initiation of RNA synthesis. To test the RdRp activity of the fusion protein, several RNA transcripts generated by in vitro transcription reactions were used as templates in the RdRp activity assay. Considering the possibility that the minor contamination of T7 RNA polymerase in the enzyme preparation might work on the residual DNA fragments containing T7 promoter and give rise to false-positive reactions in RdRp assays, SP6 RNA polymerase was used to produce the 3′-terminal fragments of both positive- and negative-strand RNAs of BaMV. As shown in Fig. 3A, the 190-nt 3′-terminal fragment of positive-strand RNA contains the 3′-untranslated region of BaMV, a poly(A) string (40 mer), and nine nonviral nucleotides in 5′→3′ order. This fragment includes the consensus sequence ACCUAA that has been shown to play an important role in the replication of clover yellow mosaic potexvirus RNA (39). The 5′-terminal fragment of positive-strand RNA has 200 bases, including 94 nt of the 5′ noncoding region and the first 106 nt of ORF1. The 3′-terminal fragment of negative-strand RNA contains 178 bases complementary to the 5′ end of positive-strand RNA. The experiment was based on the detection of incorporated labeled nucleotides in novel RNA products. The reaction mixture consisted of the RNA substrate, the four ribonucleoside triphosphates (one of which was ^{32}P labeled), and the purified fusion protein in a suitable buffer. After incubation, the nucleic

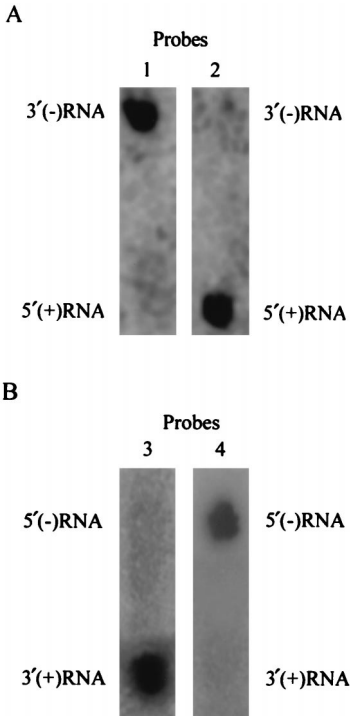

FIG. 4. Hybridization assay of products of the RdRp assay. (A) RNA fragments were cross-linked on two strips of nylon paper. Each paper had the 5′ terminus of the positive-strand RNA and the 3′ terminus of the negative-strand RNA of BaMV on each side. The paper was then incubated with the ^{32}P-labeled product of the RdRp assay in which the 3′ terminus of the negative-strand RNA had been used as the RNA substrate (lane 1). The ^{32}P-labeled 3′ terminus of the negative-strand RNA, generated by in vitro transcription, was used as the probe in another paper as a control (lane 2). (B) The 5′ terminus of the negative-strand RNA and the 3′ terminus of the positive-strand BaMV RNA were fixed on each side of a strip of nylon paper. The paper was then incubated with the ^{32}P-labeled product of the RdRp assay in which the 3′ terminus of the positive-strand RNA had been used as the RNA substrate (lane 3). The ^{32}P-labeled 3′ terminus of the positive-strand RNA, generated by in vitro transcription, was used as the probe on another RNA-fixed paper as a control (lane 4).

acids were extracted and analyzed by electrophoresis on polyacrylamide gels containing 7 M urea.

Results of the above RdRp assay are shown in Fig. 3. When the 3′-terminal fragment of either positive- or negative-strand BaMV RNA was included in the reaction mixture, a ^{32}P-labeled RNA product with a size similar to that of the RNA substrate was observed. No ^{32}P-labeled product appeared as the 5′-terminal fragment of positive-strand BaMV RNA or cucumber mosaic virus satellite RNA was used as the substrate. No reaction occurred as the enzyme was prepared from *E. coli* cells carrying plasmid pET-32a. It is also worth noting that the ^{32}P-labeled RNA products appeared only when four ribonucleoside triphosphates were present in the reaction mixture (data not shown). Since the appearance of the ^{32}P-labeled RNA products depended on the types of RNA substrate present and on the presence of four ribonucleotides, the ^{32}P-labeled products probably represent the de novo synthesizing RNA fragments due to RdRp activity rather than due to a terminal nucleotidyltransferase activity.

Dot hybridization experiments were carried out to further clarify the situation and to define the nature of the ^{32}P-labeled RNA product. If the reaction is an RdRp-catalyzed reaction, the synthesizing RNA product should be able to hybridize to the template RNA. On the other hand, the RNA product of a terminal nucleotidyltransferase should hybridize to the complementary strand of the template RNA. The 3′ terminus of negative-strand RNA and the 5′ terminus of positive-strand RNA of BaMV were fixed separately onto nylon paper by UV cross-linking. The nylon paper was subsequently incubated with the ^{32}P-labeled reaction product of an RdRp assay which had the 3′-terminal fragment of negative-strand RNA of BaMV as the substrate. The ^{32}P-labeled 3′ terminus of negative-strand RNA, generated by in vitro transcription, was used as a control. The results are shown in Fig. 4A. The ^{32}P-labeled reaction product of the RdRp assay only hybridized to the 3′-terminal fragment of negative-strand RNA, indicating that the ^{32}P-labeled product is a complementary strand of the RNA substrate in the RdRp assay. The same experiment was done to assay the nature of the RdRp product of the 3′-terminal fragment of the positive-strand RNA of BaMV (Fig. 4B). The reaction product hybridized to the 3′ terminus of positive-strand RNA but not to the 5′ terminus of negative-strand RNA. Therefore, we conclude that the 155-kDa protein of BaMV indeed exhibited RdRp activity and that such activity was preferential toward the 3′ termini of the RNA strands of BaMV.

RdRp activity of mutant and truncated proteins. The GDD motif exists in a variety of RNA-dependent RNA polymerases (18) and plays an important role in the enzymatic activity of polymerization (11, 14, 23, 24). Because of the existence of a GDD sequence, the C terminus of the 155-kDa viral protein was postulated to be a polymerase domain. To demonstrate the importance of the GDD motif in the RdRp activity, the GDD motif was changed to GAD in one mutant protein and the two aspartates were deleted from the motif in another mutant protein by PCR-based mutagenesis. Nucleotide sequencing confirmed that the desired mutations had been achieved. Both mutant proteins were expressed in *E. coli* BL21(DE3) cells and were purified by Talon metal affinity chromatography (Fig. 2). The RdRp activities of the mutant proteins were assayed to determine the importance of the GDD motif. The results show that the RdRp activity was reduced significantly in the GAD mutant protein and was abolished in the mutant protein with DD deleted (Fig. 5).

To determine whether the C terminus of the 155-kDa protein really represented a catalytic domain with polymerase activity, truncated proteins of different lengths were created and their RdRp activities were assayed. A secondary-structure

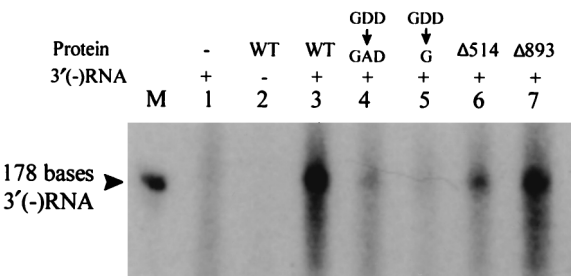

FIG. 5. Effect of mutation at the GDD motif and an N-terminal deletion of the viral protein on the RdRp activity toward the 3′ terminus of the negative-strand BaMV RNA. For lane 1, the protein prepared from *E. coli* cells harboring pET-32a was included in the assay mixture. In lane 2, wild-type fusion protein was used in the assay mixture that did not contain the RNA substrate. In lane 3, wild-type fusion protein was used in the activity assay. In lane 4, mutant protein with a GDD→GAD substitution was used in the activity assay. For lane 5, mutant protein with a GDD→G deletion was used in the activity assay. In lane 6, truncated protein containing amino acids from 514 to the C terminus was used in the activity assay, and in lane 7, truncated protein containing amino acids from 893 to the C terminus was used in the activity assay. Approximately 10 μg of each purified protein was used in the activity assay.

prediction of the 155-kDa viral protein by using the PHD mail server at the European Molecular Biology Laboratory in Heidelberg (27–29) revealed that the region between amino acids 406 and 520 and a proline-rich segment (PDPKPEPDEPLEPKAP) composed of amino acids 895 to 910 are very likely to be loop structures. If these regions really are loops, they might be the links connecting neighboring catalytic domains. Three cDNA fragments, coding for the amino acids from positions 403, 514, and 893 to the C terminus, respectively, were amplified by PCR. The DNA fragments were inserted into plasmid pET-32a, and the truncated proteins with thioredoxin, hexahistidine, and S · Tag fused at the N termini (designated Δ403, Δ514, and Δ893, respectively) were expressed in *E. coli* cells. The longest truncated protein (Δ403) was poorly expressed; the others were expressed well. The purification results for Δ514 and Δ893 by Talon chromatography are illustrated in Fig. 2. The RdRp activity remained in both Δ514 and Δ893 truncated proteins (Fig. 5), indicating that the polymerase activity resides at the C terminus of the 155-kDa viral protein.

DISCUSSION

The alphavirus-like superfamily that encompasses a large number of animal and plant virus has a 5′-capped positive-strand RNA genome. Conservation of amino acid sequence in putative methyltransferase-, helicase-, and polymerase-like domains was found within the superfamily (8, 10). These three putative catalytic domains are expressed and organized differently in different families. For instance, in potexvirus all three domains are present within a single polypeptide that is encoded by ORF1 (12, 32). In alphaviruses the methyltransferase-, helicase-, and polymerase-like domains are present in the nsP1, nsP2, and nsP4 proteins, respectively (34). How these three catalytic domains interact during the process of replication is a fundamental issue that remains to be addressed. The present study set out to identify the inherent activities and map the catalytic domains of the ORF1 product of BaMV.

To our knowledge, none of the ORF1 products of potexvirus has been expressed by heterologous expression systems. To express the 155-kDa protein of BaMV, the corresponding cDNA fragment of ORF1 was inserted into several *E. coli* expression vectors, including pMAL-c2 (NEB), pET-29a (Novagen), pET-32a (Novagen), and pMMB67HE (7), and the baculovirus expression vector pBacPAK9 (Clontech) in this study. In general, no matter what vector-host system was used, the expression of the viral protein was, with only one exception, low or even barely detected. A fair amount of the 155-kDa viral protein could be obtained in the form of a fusion protein with thioredoxin fused at its N terminus by using pET-32a vector in *E. coli* BL21(DE3) cells. Several reasons might account for this limited heterologous expression. The efficiency of the read-through translation of the 155-kDa viral protein might be poor, and the incomplete translated polypeptides were subjected to proteolytic digestion. The viral protein might not be able to fold into the correct conformation, and thus the incorrectly folded proteins may have aggregated or been subjected to proteolytic digestion. More importantly, the viral protein might be toxic to the hosts; therefore, the hosts would be killed once the viral protein was expressed. Thioredoxin might benefit the expression of the fusion protein by increasing the efficiency of translation and increasing the solubility of the protein. Therefore, under the induction condition described in Materials and Methods, a significant amount of the thioredoxin-fused protein could be obtained. Even so, many small fragments of viral protein, which might represent the immature polypeptides and/or proteolytic fragments, were found in the cell extracts (Fig. 1).

The experiments reported above showed that the product of ORF1 of BaMV is an RdRp exhibiting a preference for different RNA substrates. It catalyzed mainly the 3′-terminal fragments of positive-strand and negative-strand BaMV RNA. This preference may depend on the specific recognition of the viral protein for the tertiary structure of RNA substrates. It is noteworthy that there is no significant similarity in the nucleotide sequences of the 3′ termini of the positive-strand and the negative-strand RNAs of BaMV. Whether these two RNA fragments assume similar tertiary structures and what is the structural basis for protein recognition remain to be investigated. It is also worth noting that the 3′ terminus of negative-strand RNA appeared to be a better substrate than the 3′ terminus of positive-strand RNA, although the ^{32}P-labeled RNA products were not quantified scrupulously. This may be because the virus needs more progeny RNA, i.e., the positive-strand RNA rather than the negative-strand RNA; therefore, the synthesis of positive-strand RNA from negative-strand template should be more productive. It will be very interesting to learn how the viral protein interacts with each of the 3′ termini of BaMV RNA.

Mutational analyses also showed the important role of the GDD motif in the enzymatic activity of BaMV RdRp. The location of the GDD motif (amino acids 1228 to 1230) in the C terminus of the 155-kDa protein led to the prediction that the C terminus of the protein represents the domain with polymerase activity. To map the domain required for the RdRp activity, truncated proteins were constructed. The retention of the RdRp activity and the excellent solubility of the truncated protein Δ893 suggest that the polypeptide from amino acid 893 to the C terminus forms the polymerase domain. The reduced RdRp activity found in another truncated protein, Δ514, may result from the interference of the extra polypeptide in the folding of the polymerase domain.

In general, the capping reaction of the 5′ end of a eukaryotic mRNA is accomplished by three consecutive enzymatic reactions. The 5′ triphosphate-terminated RNA transcript is first converted to a disphosphate-terminated RNA by using RNA triphosphatase; it is then capped with GMP via a 5′-5′ triphosphate bridge by RNA guanylyltransferase and finally methylated by RNA (guanine-7-)methyltransferase. The sequence of these three catalytic reactions may vary in that the guanine is methylated before linkage to the mRNA molecule in alphavirus (1). Since the N terminus of the 155-kDa protein has several conserved amino acids characteristic of methyltransferase, the virus could possess the capping activity toward its own RNA genome. If the expressed fusion protein has the RNA triphosphatase and RNA guanylyltransferase activities, the [^{32}P]GMP moiety of [α-^{32}P]GTP would link to the 5′ end of the positive-strand RNA in a 5′-5′ linkage. No ^{32}P-labeled RNA appeared when a 5′-terminal fragment of positive-strand RNA was used as the RNA substrate in the RdRp assay (Fig. 3). No ^{32}P-labeled RNA appeared even when *S*-adenosylmethionine was included in the assay (data not shown). This result suggests that the expressed fusion protein does not have at least one of the activities of RNA triphosphatase and RNA guanylyltransferase. The fused portion at the N terminus may hinder the capping activity of the viral protein. It is also possible that plant host factors are required to assist in the capping activity. The construction of expression systems will be continued in order to produce the authentic 155-kDa viral protein for the identification of the capping activity.

ACKNOWLEDGMENTS

This work was supported by grants NSC 86-2311-B-005-022-B11 and NSC 87-2311-B-005-001-B11 from the National Science Council, Republic of China.

REFERENCES

1. **Ahola, T., and L. Kääriäinen.** 1995. Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. Proc. Natl. Acad. Sci. USA **92:**507–511.
2. **Barton, D. J., B. J. Morasco, and J. B. Flanegan.** 1996. Assays for poliovirus polymerase, $3D^{pol}$, and authentic RNA replication in HeLa S10 extracts. Methods Enzymol. **275:**35–57.
3. **Behrens, S. E., L. Tomei, and R. De Francesco.** 1996. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. EMBO J. **15:**12–22.
4. **Cho, M. W., O. C. Richards, T. M. Dmitrieva, V. Agol, and E. Ehrenfeld.** 1993. RNA duplex unwinding activity of poliovirus RNA-dependent RNA polymerase $3D^{pol}$. J. Virol. **67:**3010–3018.
5. **Cui, T., S. Sankar, and A. G. Porter.** 1993. Binding of encephalomyocarditis virus RNA polymerase to the 3′-noncoding region of the viral RNA is specific and requires the 3′-poly(A) tail. J. Biol. Chem. **268:**26093–26098.
6. **Dreher, T. W., A. L. N. Rao, and T. C. Hall.** 1989. Replication in vivo of mutant brome mosaic virus RNAs defective in aminoacylation. J. Mol. Biol. **206:**425–438.
7. **Fürste, J. P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarian, and E. Lanka.** 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tac*P expression vector. Gene **48:**119–131.
8. **Goldbach, R., O. Le Gall, and J. Wellink.** 1991. Alpha-like viruses in plants. Semin. Virol. **2:**19–25.
9. **Habili, N., and R. H. Symons.** 1989. Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerase and nucleic acid helicase. Nucleic Acids Res. **17:**9543–9555.
10. **Haseloff, J., P. Goelet, D. Zimmern, P. Ahlquist, R. Dasgupta, and P. Kaesberg.** 1984. Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organization. Proc. Natl. Acad. Sci. USA **81:**4358–4362.
11. **Hong, Y., and A. G. Hunt.** 1996. RNA polymerase activity catalyzed by a potyvirus-encoded RNA-dependent RNA polymerase. Virology **226:**146–151.

11a. **Hsu, Y.-H., C.-C. Hu, and C.-W. Lee.** Unpublished data.

11b. **Huang, C.-Y., and C.-H. Tsai.** Evolution of bamboo mosaic virus in a non-systemic host results in mutations in the helicase-like domain that cause reduced RNA accumulation. Virus Res., in press.

12. **Huisman, M. J., H. J. M. Linthorst, J. F. Bol, and B. J. C. Cornelissen.** 1988. The complete nucleotide sequence of potato virus X and its homologies at the amino acid level with various plus-stranded RNA viruses. J. Gen. Virol. **69:**1789–1798.
13. **Inokuchi, Y., and A. Hirashima.** 1987. Interference with viral infection by defective RNA replicase. J. Virol. **61:**3946–3949.
14. **Jablonski, S. A., M. Luo, and C. Morrow.** 1991. Enzymatic activity of poliovirus RNA polymerase mutants with single amino acid changes in the conserved YGDD amino acid motif. J. Virol. **65:**4565–4572.
15. **Jablonski, S. A., and C. Morrow.** 1995. Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal ion requirements for activity. J. Virol. **69:**1532–1539.
16. **Kadaré, G., and A.-L. Haenni.** 1997. Virus-encoded RNA helicases. J. Virol. **71:**2583–2590.
17. **Kao, C. C., and J.-H. Sun.** 1996. Initiation of minus-strand RNA synthesis by the brome mosaic virus RNA-dependent RNA polymerase: use of oligoribonucleotide primers. J. Virol. **70:**6826–6830.
18. **Koonin, E. V.** 1991. The phylogeny of RNA-dependent RNA polymerase of positive-strand RNA viruses. J. Gen. Virol. **72:**2197–2206.
19. **Koonin, E. V., and V. V. Dolja.** 1993. Evolution and taxonomy of positive-strand RNA viruses: implication of comparative analysis of amino acid sequence. Biochem. Mol. Biol. **28:**375–430.
20. **Kroner, P., D. Richards, P. Traynor, and P. Ahlquist.** 1989. Defined mutations in a small region of the brome mosaic virus 2a gene cause diverse temperature-sensitive RNA replication phenotypes. J. Virol. **63:**5302–5309.
21. **Lin, N.-S., F.-Z. Lin, T.-Y. Huang, and Y.-H. Hsu.** 1992. Genome properties of bamboo mosaic virus. Phytopathology **82:**731–734.
22. **Lin, N.-S., B.-Y. Lin, N.-W. Lo, C.-C. Hu, T.-Y. Chow, and Y.-H. Hsu.** 1994. Nucleotide sequence of the genomic RNA of bamboo mosaic potexvirus. J. Gen. Virol. **75:**2513–2518.
23. **Lohmann, V., F. Körner, U. Herian, and R. Bartenschlager.** 1997. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. J. Virol. **71:**8416–8428.
24. **Longstaff, M., G. Brigneti, F. Boccard, S. Chapman, and D. Baulcombe.** 1993. Extreme resistance to potato virus X infection in plants expressing a modified component of the putative viral replicase. EMBO J. **12:**379–386.
25. **Osman, T. A. M., and K. W. Buck.** 1996. Complete replication in vitro of tobacco mosaic virus RNA by a template-dependent, membrane-bound RNA polymerase. J. Virol. **70:**6227–6234.
26. **Pilipenko, E. V., S. V. Maslova, A. N. Sinyakov, and V. I. Agol.** 1992. Towards identification of *cis*-acting elements involved in the replication of enterovirus and rhinovirus RNAs: a proposal for the existence of tRNA-like terminal structures. Nucleic Acids Res. **20:**1739–1745.
27. **Rost, B., and C. Sander.** 1993. Prediction of protein structure at better than 70% accuracy. J. Mol. Biol. **232:**584–599.
28. **Rost, B., C. Sander, and R. Schneider.** 1994. PHD—an automatic mail server for protein secondary structure prediction. Comput. Appl. Biosci. **10:**53–60.
29. **Rost, B., and C. Sander.** 1994. Combining evolutionary information and neural networks to predict protein secondary structure. Proteins **19:**55–72.
30. **Rozanov, M. N., E. V. Koonin, and A. E. Gorbalenya.** 1992. Conservation of the putative methyltransferase domain: a hallmark of the "Sindbis-like" supergroup of positive-strand RNA viruses. J. Gen. Virol. **73:**2129–2134.
31. **Siegel, R. W., S. Adkins, and C. C. Kao.** 1997. Sequence-specific recognition of a subgenomic RNA promoter by a viral RNA polymerase. Proc. Natl. Acad. Sci. USA **94:**11238–11243.
32. **Skryabin, K. G., S. Y. Morozov, A. S. Kraev, M. N. Rozanov, B. K. Chernov, L. I. Lukasheva, and J. G. Atabekov.** 1988. Conserved and variable elements in RNA genomes of potexviruses. FEBS Lett. **240:**33–44.
33. **Smirnyagina, E., N.-S. Lin, and P. Ahlquist.** 1996. The polymerase-like core of brome mosaic virus 2a protein, lacking a region interacting with viral 1a protein in vitro, maintains activity and 1a selectivity in RNA replication. J. Virol. **70:**4729–4736.
34. **Strauss, J. H., and E. G. Strauss.** 1994. The alphaviruses: gene expression, replication, and evolution. Microbiol. Rev. **58:**491–562.
35. **Sun, J.-H., S. Adkins, G. Faurote, and C. C. Kao.** 1996. Initiation of (−)-strand RNA synthesis catalyzed by the BMV RNA-dependent RNA polymerase: synthesis of oligonucleotides. Virology **226:**1–12.
36. **Tabor, S., H. E. Huber, and C. C. Richardson.** 1987. *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. J. Biol. Chem. **262:**16212–16223.
37. **Tsai, C.-H., and T. W. Dreher.** 1991. Turnip yellow mosaic virus RNAs with anticodon loop substitutions that result in decreased valylation fail to replicate efficiently. J. Virol. **65:**3060–3067.
38. **Vázquez, A. L., J. M. M. Alonso, R. Casais, J. A. Boga, and F. Parra.** 1998. Expression of enzymatically active rabbit hemorrhagic disease virus RNA-dependent RNA polymerase in *Escherichia coli*. J. Virol. **72:**2999–3004.
39. **White, K. A., J. B. Bancroft, and G. A. Mackie.** 1992. Mutagenesis of a hexanucleotide sequence conserved in potexvirus RNAs. Virology **189:**817–820.