Non-AUG Translation Initiation in a Plant RNA Virus: a Forty-Amino-Acid Extension Is Added to the N Terminus of the Soil-Borne Wheat Mosaic Virus Capsid Protein

YUKIO SHIRAKO

Asian Center for Bioresources and Environmental Sciences, University of Tokyo, Bunkyo-ku, Tokyo, Japan; Division of Biology, California Institute of Technology, Pasadena, California; and Center for Agricultural Molecular Biology, Cook College, Rutgers University, New Brunswick, New Jersey

Received 2 June 1997/Accepted 25 October 1997

RNA 2 of soil-borne wheat mosaic virus (SBWMV), the type species of the genus Furovirus, encodes a protein previously hypothesized to be initiated at an in-frame non-AUG codon upstream of the AUG initiation codon (nucleotide positions 334 to 336) for the 19-kDa capsid protein. Site-directed mutagenesis and in vitro transcription and translation analysis indicated that CUG (nucleotides 214 to 216) is the initiation codon for a protein with a calculated molecular mass of 25 kDa composed of a 40-amino-acid extension to the N terminus of the 19-kDa capsid protein. A stable deletion mutant, which was isolated after extensive passages of a wild-type SBWMV, contained a mixture of two deleted RNA 2's, only one of which coded for the 25-kDa protein. The amino acid sequence of the N-terminal extension was moderately conserved and the CUG initiation codon was preserved among three SBWMV isolates from Japan and the United States. This CUG codon sequence conservation, as well as the retention of expression of the 25-kDa protein in the stable deletion mutant, suggests that the 25-kDa protein is functional in the life cycle of SBWMV. This is the first report of a non-AUG translation initiation in a plant RNA virus genome.

Soil-borne wheat mosaic virus (SBWMV) is the type species of the genus Furovirus (8). The group is characterized by transmission through plasmодиophoraceous fungus in soil and a divided plus-sense RNA genome, with each RNA individually encapsidated into rigid rod-shaped particles (7, 35). The SBWMV genome consists of two RNA species, RNA 1 (7,099 nucleotides) and RNA 2 (3,593 nucleotides), both of which are divided plus-sense RNA genomes, with each RNA individually encapsidated into rigid rod-shaped particles (7, 35). The SBWMV isolation of the genus Furovirus encodes a protein previously hypothesized to be initiated at an in-frame non-AUG codon upstream of the AUG initiation codon (nucleotide positions 334 to 336) for the 19-kDa capsid protein. Site-directed mutagenesis and in vitro transcription and translation analysis indicated that CUG (nucleotides 214 to 216) is the initiation codon for a protein with a calculated molecular mass of 25 kDa composed of a 40-amino-acid extension to the N terminus of the 19-kDa capsid protein. A stable deletion mutant, which was isolated after extensive passages of a wild-type SBWMV, contained a mixture of two deleted RNA 2's, only one of which coded for the 25-kDa protein. The amino acid sequence of the N-terminal extension was moderately conserved and the CUG initiation codon was preserved among three SBWMV isolates from Japan and the United States. This CUG codon sequence conservation, as well as the retention of expression of the 25-kDa protein in the stable deletion mutant, suggests that the 25-kDa protein is functional in the life cycle of SBWMV. This is the first report of a non-AUG translation initiation in a plant RNA virus genome.

In vitro mutagenesis of CUG and GUG codons upstream from the capsid protein gene. A full-length SBWMV RNA 2 cDNA of the Nebraska isolate (1988) was cloned downstream...
from the bacteriophage T7 RNA polymerase promoter in pGEM3Z (Promega) as follows. RNA 2 cDNA was synthesized with avian myeloblastosis virus reverse transcriptase (Life Sciences) from the 3' end with the primer 5'-TGCTCTAGATGGCCGGATAAACCCCTCGG-3', in which the 6 underlined bases identify an XbaI site for linearization of the template DNA and the 20 bases in boldface type are complementary to the 3'-terminal sequence of SBWMV RNA 2. The 3.6-kb cDNA was made double stranded with Escherichia coli DNA polymerase I (Pharmacia) and RNase H (Pharmacia), flush-ended with T4 DNA polymerase (Pharmacia), digested with XbaI, and ligated into Smal-XbaI-restricted pGEM3Z. For cloning the 5'-terminal region, including an upstream T7 RNA polymerase promoter sequence, cDNA was synthesized with a minus-sense primer annealing to nucleotides 113 to 131 and amplified by PCR with T7 DNA polymerase (Promega) by use of an additional plus-sense, upstream primer, 5'-GGCATTAATACGACTCACTATA-3', in which the 6 underlined bases identify an XbaI site, the 17 italicized bases are the T7 RNA polymerase promoter, and the 23 boldface bases are a unique NdeI site. The 17 italicized bases are the 3' terminal region of the SBMV RNA 2. The PCR products were digested with NdeI and HpaI at nucleotide 26 and ligated into an NdeI-HpaI cut vector DNA which contains the cDNA insert ranging from the HpaI site to the 3' terminus with an extra XbaI site. The nucleotide sequence of the resulting full-length cDNA clone, pSW2.4, was determined by dideoxy sequencing and confirmed to be identical to the sequence of the 1981 Nebraska isolate (34).

Figure 1A shows the genome organization of SBWMV RNA 2. Figure 1B shows the nucleotide and amino acid sequences between nucleotides 184 and 345, including the 5'-terminal region of the capsid protein gene. Between nucleotides 187 and 256, there are three CUG triplets at positions 214 to 216, 250 to 252, and 256 to 258, one GUG triplet at position 244 to 246, and no ACG or AUU codon (Fig. 1B). Of the three CUG codons, only the one at positions 214 to 216 is in the preferred context, i.e., G at both −3 and +4 positions as well as A at +5 and U at +6. Given this, the U at position 215 was replaced with A, changing the potential CUG start codon to CAG, by in vitro mutagenesis (24). Similarly, U at position 245 was replaced with A or C to replace the GUG codon at position 244 to 246 with GAG or GCG, respectively. Mutant plasmid DNA was linearized with XbaI and transcribed in vitro with T7 RNA polymerase (Pharmacia) in the presence of a cap analog (New England Biolabs). In vitro transcripts were translated in vitro in rabbit reticulocyte lysates (Promega) in the presence of [35S]methionine (1,200 Ci/mol; Amersham). Translation products were separated in an SDS-12.5% polyacrylamide gel and visualized by fluorography using sodium salicylate (9).

Figure 2 shows in vitro translation products directed by in vitro transcripts from wild-type and different mutant constructs. From the wild-type transcripts, the 19-kDa capsid protein, 28K protein, and 84-kDa capsid readthrough protein (84-kDa CP/RT) were expressed (lane 1). When nucleotide 215 was changed from a U to an A, resulting in replacement of the CUG with CAG, the capsid and 84-kDa readthrough proteins were produced but 28K protein was no longer expressed (lane 2). By contrast, when U at position 245 was substituted with A or C, resulting in replacement of the GUG with GAG or GCG, respectively, the 28K protein was expressed in addition to the capsid and 84-kDa readthrough proteins (lanes 3 and 4). Two or three extra, faint products detected between the 28K protein and the 84-kDa readthrough protein probably initiated internally at AUG codons downstream from the opal (UGA) termination codon for the capsid protein gene as previously described (34). Accumulation of these additional translation products varied among different translations and depended on batches of reticulocyte lysates as well as in vitro transcripts. This result indicates that the CUG codon at position 214 to 216 is the initiation codon for the 28K polypeptide, with the calculated molecular mass of 25 kDa. Hereafter the 28K protein is referred to as the 25-kDa protein.
FIG. 2. [35S]methionine-labelled in vitro translation products in rabbit reticulocyte lysates analyzed by SDS–12.5% PAGE and fluorography. Translation was directed with in vitro transcripts from the wild-type full-length cDNA construct pSW2.4 (lane 1) and from mutant constructs in the pSW2.4 background with T215→A (lane 2), T245→A (lane 3), and T245→C (lane 4) substitutions. The 25-kDa N-terminal extension and capsid protein marker (N-ext/CP) in the right margin indicates the position of the 28K protein. RT, readthrough protein.

Synthesis of the 25-kDa protein from the Lab 1 deletion mutant RNA 2. SBWMV generates spontaneous deletion mutants when wild-type virus is maintained or passed in wheat plants for prolonged periods at around 17°C (10–12, 17, 31–33). A stable deletion mutant, the Lab 1 isolate, was isolated from a wild-type infection originating in an infested field in Nebraska in 1975 after successive mechanical transfers in Triticum aestivum L., cultivar Michigan Amber) for 2 years in a growth cabinet (5, 6). Previously, it was shown that the 25-kDa protein is produced by the Lab 1 isolate in amounts similar to those of the wild-type virus both in vivo and in vitro (17). To verify the 25-kDa protein-encoding sequence and to identify deletion sites in the Lab 1 mutant genome, the complete nucleotide sequence of the Lab 1 RNA 2 was determined from two cDNA clones, p706 and p715, and from viral RNA for the 5'-terminal region as previously described (34) (Fig. 3A). In the 5'-terminal region upstream from the capsid protein gene, there were two single nucleotide deletions, a U from a U trimer (wild-type nucleotides 144 to 146) and another U from a U dimer (nucleotides 332 to 333), the latter of which was located immediately upstream of the capsid protein gene initiation codon. When another 10 independent cDNA clones were further examined for these two deletions, a U was deleted from nucleotides 144 to 146 in all 10 clones. Three clones retained the U dimer at nucleotides 332 to 333, whereas 7 clones had the U deletion, as was shown in clones p706 and p715. On the other hand, in the capsid-readthrough region, there were a 108-nucleotide (wild-type nucleotide positions 866 to 973), in-frame deletion, beginning one C residue after the capsid protein UGA termination codon, and a 1,058-nucleotide deletion (wild-type nucleotide positions 1469 to 2526) in the C-terminal region, resulting in truncation of the wild-type 84-kDa readthrough protein to 39 kDa (Fig. 3A). Thus, in the deleted capsid-readthrough polypeptide, only the wild-type amino acid at positions 214 to 378 (the position 1 is the capsid N terminus) was fused to the C terminus of the capsid protein. There were no further nucleotide deletions in the 3'-terminal region.

The Lab 1 cDNA clone p706, which had deletions of the two single U’s, was transcribed in vitro with the upstream T7 promoter, and the resulting transcripts were translated in vitro in rabbit reticulocyte lysates. Compared with the wild-type translation products (Fig. 3B, lane 1), the p706 transcripts directed synthesis of only the 19-kDa capsid protein and the capsid readthrough protein with a calculated molecular mass of 39 kDa but not the 25-kDa polypeptide (lane 2). Deletion of the U immediately upstream of the AUG capsid initiation codon causes a frameshift of the translation initiated at the CUG codon, resulting in termination of translation at the third amino acid after the deletion. From these results, it appears that the Lab 1 isolate is a mixture of two mutants, only one of which codes for the 25-kDa protein.

Conservation of the amino acid sequences in the N-terminal extension. To examine conservation of the amino acid sequences in the N-terminal extension among isolates of SBWMV from different geographic regions, nucleotide sequences of the 5'-terminal 0.9 kb of RNA 2 were determined from the wild-type, nondeleted forms of an Illinois isolate from 1990 (a kind gift from Adriana Hewings, University of Illinois, Urbana) and a Japanese isolate, JT, from 1982 (33). Figure 4 shows nucleotide and amino acid sequence alignments of the three isolates in the 5'-terminal untranslated region (5’ UTR, panel A), the 120-nucleotide N-terminal extension region (panel B), and the capsid protein gene (panel C). Table 1 shows percent identities among the three isolates in each region at the nucleotide and amino acid levels. In the three
regions combined, the Illinois isolate was 89% identical to the Nebraska isolate at the nucleotide level while there was 68% identity between the Japanese isolates and either of the two U.S. isolates. In both the Illinois and Japanese isolates, there was a CUG codon 120 nucleotides upstream from the capsid protein AUG initiation codon, similar to the Nebraska isolate. Significant amino acid sequence identities were found in the N-terminal extension region among the three isolates (Table 1). The N-terminal extension was rich in charged amino acids, particularly in the N-terminal half, whereas the C-terminal region preceding the capsid protein possessed clusters of hydrophobic amino acids in all three isolates (Fig. 4B). On the other hand, the capsid protein was more conserved than the N-terminal extension at both nucleotide and amino acid levels among the three isolates. The capsid proteins of the two U.S. isolates were identical at the amino acid level, while that of the Japanese isolate was 82% identical to those of the two U.S. isolates (Fig. 4C and Table 1).

The 25-kDa protein is possibly functional in the SBWMV life cycle. Previously, it was shown that the 25-kDa protein...
TABLE 1. Percent identities at the nucleotide and amino acid levels in the 5′-untranslated region, the N-terminal extension region, and the capsid protein-coding region among Nebraska (NE), Illinois (IL), and Japanese (JT) isolates of SBWMV.

<table>
<thead>
<tr>
<th>Region</th>
<th>% Identity at the nucleotide level</th>
<th>amino acid level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NE vs IL</td>
<td>NE vs JT</td>
</tr>
<tr>
<td>5′-Untranslated region*</td>
<td>96</td>
<td>59</td>
</tr>
<tr>
<td>N-terminal extensionb</td>
<td>87/90</td>
<td>60/65</td>
</tr>
<tr>
<td>Capsid protein</td>
<td>87/100</td>
<td>74/82</td>
</tr>
</tbody>
</table>

* Nucleotide positions 1 to 213 of NE and IL and positions 1 to 209 of JT were compared. Gaps were included to maximize the alignment as shown in Fig. 4A. 

b Nucleotide positions 214 to 335 of NE and IL and positions 210 to 329 of JT were compared.

and equivalent proteins were produced not only by wild-type SBWMV but also by extensively passaged stable deletion mutants (17, 33). Although the Nebraska Lab 1 isolate used in this study was a mixture of two mutants based on the deletion of a U at positions 331 to 332, it appeared that there was a selection pressure preventing the deletion of the U in the RNA 2 population. It is not likely that the U is required for efficient translation initiation of the capsid protein gene because the deletion mutant RNA 2 (clone p706) produced a similar amount of capsid protein compared with the wild type (Fig. 3B, lanes 1 and 2), and the context surrounding the capsid initiation codon still retained the preferred consensus sequence with a purine at the +3 position even after the deletion. Rather, the translated 25-kDa protein itself is probably required for replication, movement, and/or spreading of the virus. The significant amino acid sequence conservation in the N-terminal extension among the three isolates from different geographic regions also suggests that the 25-kDa protein is functional in the SBWMV life cycle. Similarities in the N-terminal extension at the amino acid level among the three isolates were fewer than those in the capsid protein, and the N-terminal half of the extension is highly charged, suggesting that the N-terminal extension is not as highly structured as the capsid protein.

It is noteworthy that contexts surrounding the CUG codon as well as the capsid initiation codon in the Japanese isolate do not have the preferred sequence (3), with a purine at the +3 position even after the deletion. Rather, the translated 25-kDa protein itself is probably required for replication, movement, and/or spreading of the virus. The significant amino acid sequence conservation in the N-terminal extension among the three isolates from different geographic regions also suggests that the 25-kDa protein is functional in the SBWMV life cycle. Similarities in the N-terminal extension at the amino acid level among the three isolates were fewer than those in the capsid protein, and the N-terminal half of the extension is highly charged, suggesting that the N-terminal extension is not as highly structured as the capsid protein.

I am grateful to James Strauss of the California Institute of Technology and Peter Day of Rutgers University for their support in this study and to Myron Brakke for critical reading of the manuscript.

REFERENCES


Downloaded from http://jvi.asm.org/ on October 20, 2017 by guest