The complete nucleotide sequence of RNA 1 of a German isolate of barley yellow mosaic virus and its comparison with a Japanese isolate

E. Peerenboom,* M. Pröls, J. Schell, H.-H. Steinbiss and A. D. Davidson

Max-Planck-Institut für Züchtungsforschung, Abteilung Genetische Grundlagen der Pflanzenzüchtung, Carl-von-Linné Weg 10, D-5000 Köln 30, Germany

The nucleotide sequence of RNA 1 of a German isolate of barley yellow mosaic virus has been determined and compared with a Japanese isolate of the same virus. The sequence identity is 93.6% at the nucleotide level and 96% at the amino acid level. Similar values have been found for the polyproteins of the RNA 2 of both isolates (95%). Both isolates show an RNA 1-encoded protein arrangement similar to that of potyviruses such as tobacco etch virus. In contrast, the polyproteins of the small RNAs (RNA 2) do not show such a similarity to the polyproteins of other potyviruses. However, there is a striking difference between the two isolates in the generally highly conserved active site of the RNA-dependent RNA polymerase. The German isolate exactly matches the consensus sequences for previously described potyviral RNA-dependent RNA polymerases, whereas the Japanese isolate does not.

Barley yellow mosaic virus (BaYMV) together with some other filamentous, soil-borne viruses was originally classified as a possible member of the potyvirus group (Matthews, 1982). This classification was based on their particle morphology and the induction of characteristic cytoplasmic cylindrical inclusion bodies in the cereal hosts (Hibino et al., 1981; Ebrahim-Nesbat & Zerlik, 1984; Huth et al., 1984). However, in contrast to potyviruses this group of viruses has a genome consisting of two RNA species (Huth, 1988; Usugi et al., 1989) and is transmitted by the fungus Polymyxa graminis (Adams et al., 1989; Brunt, 1989). In addition, most of these fungus-transmitted viruses are serologically related to each other but most do not show any serological relationship to potyviruses transmitted by other vectors (Huth et al., 1984; Ehlers & Paul, 1986; Adams et al., 1987; Usugi et al., 1989). More recently it has been proposed that this potyvirus subgroup should form the basis of a new genus bymovirus of the family Potyviridae, with BaYMV, the best characterized of these viruses, as the type member (Usugi et al., 1989; Kashiwazaki et al., 1990; Barnett et al., 1991).

Recently, RNA 1 and RNA 2 of a Japanese isolate (BaYMV-J) (Kashiwazaki et al., 1990, 1991) and RNA 2 of a German isolate (BaYMV-G) (Davidson et al., 1991) of BaYMV have been cloned and sequenced. This has facilitated a comparison of BaYMV with potyviruses at the nucleotide sequence level. The genome organization and possible protein maturation and functions of BaYMV show a close but nevertheless distinct relationship to that of other potyviruses. In this paper we present the nucleotide sequence of RNA 1 of BaYMV-G and compare it with that of BaYMV-J. This completes the sequencing of the German BaYMV isolate and gives further insight into the relatedness of different BaYMV isolates and their relationship to the potyvirus group in general.

RNA isolated from a German field isolate of BaYMV was used as previously described (Pröls et al., 1990) to synthesize the cDNA clone pLY29, which corresponds to nucleotides (nts) 1423 to 7643 of BaYMV RNA 1. DNA sequence information derived from this clone was used to design the oligonucleotide primers 992 and 1276, complementary to nts 1453 to 1472 and 1488 to 1515 of RNA 1. The amplified cDNA population was cloned and then screened by restriction enzyme analysis. The largest cDNA inserts (approx. 1.4 kb) were sequenced at the 5' end and the data used to design the primer 2109, complementary to nts 116 to 137 of RNA 1. Primer 2109 was then used for primer extension analysis and subsequent PCR-mediated cDNA cloning (Davidson et al., 1991) of the 5' end of RNA 1. cDNA clones corresponding in size to the primer extension products.
were sequenced and used to design the primer 5' GGCGGCCCCTGCAATAGCAGACTCATTAGA- AAAATAAAACCCCTAACC 3' containing a NotI restriction site, the bacteriophage T7 promoter and the first 23 nts of RNA 1. This primer was used in conjunction with primer 1276 for PCR-mediated cDNA cloning (Sambrook et al., 1989) of the first 1515 nts of RNA 1. cDNA fragments of the expected size were cut with NotI and NsiI, isolated and cloned into the corresponding sites of plY29 to give the cDNA clone plY29T7 which corresponds to the complete genome of RNA 1.

The dideoxynucleotide chain termination method (Sanger et al., 1977) was used to sequence both strands of subclones derived from plY29 using both restriction enzyme fragments and exonuclease III-generated nested enzyme fragments and exonuclease III-generated nested fragments isolated from independent PCR reactions. As the error rate of the PCR is relatively high (Erlich et al., 1991), three plY29T7 clones synthesized using PCR primers 2254 (complementary to nts 179 to 180) which was predicted not to hybridize to RNA 2. This confirmed the determined sequence of RNA 1. Although each of the three PCR cDNA clones differed by a one nt substitution, attributable to either RNA population heterogeneity or the infidelity of the PCR reaction, sequencing of the PCR-generated fragments in three independent clones allowing the compilation of the 7643 nucleotide sequence [excluding the poly(A) tail] corresponding to RNA 1 of BaYMV-G shown in Fig. 1.

The polyproteins of BaYMV-G and BaYMV-J share sequence for translation initiation (CUUAUGG) differs from that of BaYMV-J (CCUAUGG) and may be less favourable for translation initiation (Kozak, 1986; Liitcke et al., 1987). In addition, the sequence CCUAUGG which contains a second in-frame AUG codon is repeated in both isolates at positions 225 to 231 (BaYMV-G). Thus the point of translation initiation could differ in the two isolates. If this is the case, the first 18 amino acids encoded by BaYMV-G shown in Fig. 1 would be absent. This awaits further investigation and in this paper the ORF discussed is that shown in Fig. 1. Computer translation of the nucleotide sequence in both viral positive and negative strands predicts only one long, open reading frame (ORF) which, by comparison with BaYMV-J, probably commences at the first AUG codon (position 174) and terminates with a UAA codon at position 7410. However, the consensus sequence for translation initiation (CUUAUGG) differs from that of BaYMV-J (CCUAUGG) and may be less favourable for translation initiation (Kozak, 1986; Liitcke et al., 1987). In addition, the sequence CCUAUGG which contains a second in-frame AUG codon is repeated in both isolates at positions 225 to 231 (BaYMV-G). Thus the point of translation initiation could differ in the two isolates. If this is the case, the first 18 amino acids encoded by BaYMV-G shown in Fig. 1 would be absent. This awaits further investigation and in this paper the ORF discussed is that shown in Fig. 1. Thus the ORF codes for a protein of 2412 amino acids (Mr 270873), two amino acids larger than the protein (Mr 270855) with two mismatches could not be dismissed. Thus further extension and cDNA cloning steps were performed using the primer 2254 (complementary to nts 179 to 178) which was predicted not to hybridize to RNA 2. Although each of the three PCR cDNA clones differed by a one nt substitution, attributable to either RNA population heterogeneity or the infidelity of the PCR reaction, sequencing of the PCR-generated fragments in three independent clones allowing the compilation of the 7643 nucleotide sequence [excluding the poly(A) tail] corresponding to RNA 1 of BaYMV-G shown in Fig. 1.

The dideoxynucleotide chain termination method (Sanger et al., 1977) was used to sequence both strands of subclones derived from plY29 using both restriction enzyme fragments and exonuclease III-generated nested deletions (Henikoff, 1987). Certain regions of plY29 were also sequenced using the specific internal primers. As the error rate of the PCR is relatively high (Erlich et al., 1991), three plY29T7 clones synthesized using PCR fragments isolated from independent PCR reactions were sequenced using specific primers in the region of nts 1 to 1515 of RNA 1. There was a one bp difference in two PCR clones, each in a different location. The consensus of these sequences was used to compile the sequence of RNA 1. Nucleotide sequence data were assembled and analysed using a VAX 220 computer and the software package of Devereux et al. (1984).

Initially, primer extension and sequence analysis of the 5' region of RNA 1 led to the detection of a putative 5' end which had high identity to that previously detected for RNA 2 of BaYMV (Davidson et al., 1991). As total BaYMV RNA was used for these experiments the possibility that the primer 2160 could bind to RNA 2 with two mismatches could not be dismissed. Thus further extension and cDNA cloning steps were performed using the primer 2254 (complementary to nts 179 to 178) which was predicted not to hybridize to RNA 2. Although each of the three PCR cDNA clones differed by a one nt substitution, attributable to either RNA population heterogeneity or the infidelity of the PCR reaction, sequencing of the PCR-generated fragments in three independent clones allowing the compilation of the 7643 nucleotide sequence [excluding the poly(A) tail] corresponding to RNA 1 of BaYMV-G shown in Fig. 1.
the amino acid positions in the polyproteins.

Identical amino acids are marked by asterisks. The numbering refers to the amino acid positions in the polyproteins.

Fig. 2. Comparison of the amino acid motif in the active site of the RNA-dependent RNA polymerase. The top line shows the common arrangement and function of the proteins as was found (Argos, 1984; Domier, 1987). Whereas BaYMV-J coat proteins have only three amino acids conserved (Fig. 2), the German isolate Japanese isolate

Table 1. Differences at the sequence level in RNA 1 of BaYMV-J and BaYMV-G

<table>
<thead>
<tr>
<th></th>
<th>German isolate</th>
<th>Japanese isolate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length*</td>
<td>7643</td>
<td>7631</td>
</tr>
<tr>
<td>Longest ORF*</td>
<td>7236</td>
<td>7230</td>
</tr>
<tr>
<td>5' non-translated region</td>
<td>174</td>
<td>172</td>
</tr>
<tr>
<td>3' non-translated region</td>
<td>235</td>
<td>231</td>
</tr>
<tr>
<td>Consensus sequence for translation initiation†</td>
<td>GCUCUUAUGGAG</td>
<td>GCUCUUAUGGAG</td>
</tr>
<tr>
<td>Length of encoded polyprotein†</td>
<td>2412</td>
<td>2410</td>
</tr>
<tr>
<td>Mr of polyprotein</td>
<td>270873</td>
<td>270755</td>
</tr>
</tbody>
</table>

* Length in nts.
† The difference in the consensus sequence for translation initiation is underlined.
‡ Length in amino acids.

calculated for the polyprotein of RNA 2 with an amino acid identity of 94%. These differences could be due to the high natural variability of RNA viruses which lack an RNA repair system (Holland et al., 1982; Reanney, 1984). Computer analysis allowed prediction of the same arrangement and function of the proteins as was found for different isolates of a potyvirus (Shukla & Ward, 1988). Other differences between BaYMV-G and BaYMV-J at the sequence level are shown in Table 1.

References


(Received 16 September 1991; Accepted 3 January 1992)