European Journal of Plant Pathology **100**: 233–241, 1994.

The complete nucleotide sequence of RNA2 of barley mild mosaic virus (BaMMV)

ULRICH TIMPE and THOMAS KÜHNE
Federal Centre for Breeding Research on Cultivated Plants, Institute for Resistance Research, Theodor-Roemer-Weg 4, D-06449 Aschersleben, Germany

Accepted 20 April 1994

Key words: BaMMV, barley mild mosaic virus, *bymovirus*, nucleotide sequence, *potyviridae*, RNA2

Abstract. Barley mild mosaic virus is a member of the *Bymoviruses*, a genus of the family *Potyviridae*. The virus consists of two types of flexuous rod-shaped particles. Each of them contains one single-stranded polyadenylated RNA in plus orientation of approximately 7.6 kb (RNA1) and 3.6 kb (RNA2). Complementary DNAs of both RNAs have been synthesised and cloned. The nucleotide sequence of RNA2 has been determined. It is 3524 nucleotides in length, excluding the 3′ poly(A) tail, and contains one large open reading frame (2679 nts), coding for a polyprotein of approximately 98 kDa. There are indications that a putative proteolytic activity in the N-terminal part can cleave the polyprotein autocatalytically into a 25 kDa protein (putative proteinase) and a 73 kDa polypeptide of unknown function.

Introduction

Barley mild mosaic virus (BaMMV) has been identified as one of the agents causing the yellow mosaic disease of winter barley. In nature this virus is transmitted by the soil-borne plasmodiophoid fungus *Polymyxa graminis* Ledingham [Adams et al., 1989]. The particles remain infective in resting spores of the fungus for many years. BaMMV possesses slightly flexuous rod-shaped particles with a bimodal length distribution of 270 to 289 nm and 568 to 600 nm [Huth et al., 1984]. Both genomic RNA molecules are single-stranded and have a poly(A) tail at their 3′ terminus. At the 5′ terminus they are expected to carry a genome-linked protein [Usugi et al., 1989]. RNA1 and RNA2 have molecular weights of approximately 2.5 MDa and 1.2 MDa, respectively. They are encapsidated by one type of capsid protein of about 35 kDa [Ehlers and Paul, 1986].

In winter barley BaMMV frequently occurs in mixed infections with barley yellow mosaic virus (BaYMV), which exhibits similar morphological and biological properties but can be differentiated serologically [Huth, 1986; Huth and Adams, 1990]. Recently, complete sequences of both RNAs of a Japanese and a German isolate of BaYMV have been published [Kashiwazaki et al., 1990, 1991; Davidson et al., 1991; Peerenboom et al., 1992], and the 3′-terminal part of BaMMV RNA1 has also been sequenced [Kashiwazaki et al., 1992; Schlichter et al., 1993].

Here we present the complete nucleotide sequence of RNA2 of a German isolate of BaMMV.

Materials and methods

Virus purification and RNA preparation

The isolate BaMMV-ASL1 originating from a field at Aschersleben (Germany) was propagated in a growth chamber on mechanically inoculated barley plants (cv. Erfa) and purified using the method of Huth et al. [1984]. Viral RNA was isolated from particles according to Kashiwazaki et al. [1989] with slight modifications (0.5% SDS, additional chloroform extraction step) and omitting the final oligo(dT) affinity chromatography.

cDNA synthesis and cloning

A mixture of both, RNA1 and RNA2 was transcribed into cDNA according to Gubler and Hoffman [1983] by using oligo(dT)primers. After synthesis of second strand DNA, treatment of the product with T4 DNA polymerase was made to generate blunt ends. Eco RI adapters were linked to the termini and the DNA fragments were ligated into bacteriophage λgt 11 arms and cloned in *E. coli* strain Y 1090 according to the instructions for the cDNA Cloning System Plus (Amersham, Braunschweig, Germany).

Two clones, L12 (1.2 kbp) and L27 (3.6 kbp) were identified by hybridisation from the genomic library which corresponded with BaMMV RNA2. After Eco RI digestion, the complete L12 insert and two L27 subfragments (0.8 and 2.8 kbp) were ligated into pGEM-3Zf(–) vector (Promega, Heidelberg, Germany), cloned and detected in both orientations as pG12(+) and pG12(–), pG2708(+) and pG2708(–), pG2728(+) and pG2728(–). To clone the complete insert of L27, a Bam HI fragment was ligated into the plasmid vector to give pG27. The extreme 5′ terminus of RNA2 was cloned using the 5′-AmpliFINDER RACE kit (Clontech, Palo Alto, USA) and the virus specific cDNA primer 5′-GTGTGATATCAGCAGTGTCG-3′ and the nested PCR primer 5′-ATGGAACTGCAGTGATCGGC-3′ according to manufacturer's instructions. The 700 bp PCR product was cloned to give pG5′.

Sequencing

Both ssDNA and dsDNA sequencing was performed by the chain termination method [Sanger et al., 1977]. The clones, described above, but not pG27, and a set of restriction deleted subclones were used to determine overlapping sequences of both cDNA strands. In order to confirm the internal Eco RI single restriction site of pG27 a Hinc II-deleted subclone was also sequenced.

Results and discussion

The RNA2 of the isolate BaMMV-ASL1 comprises 3524 nucleotides (nts), excluding the 3′ poly(A) tail (Fig. 1). The cloning of the extreme 5′ terminus was performed with both a RNA preparation which had been stored for almost 3 years at −80 °C and a freshly prepared RNA. Following RACE PCR a number of independent clones, representing both RNA preparations was produced and sequenced. All preparations extended the 5′ terminus of pG2708 by 14 nts and were identical in sequence. This confirmed that the extreme 5′ terminus had been reached. However the possible presence of a covalently linked protein on the viral RNA might have hindered complete cDNA synthesis and precluded determination of the terminal nt.

Clone pG12 contains a track of about 80 A residues, indicating that the 3′ terminus is covered by this clone.

The 140 nts long 5′ untranslated region (UTR) is very rich in A and U, but poor in G similarly to other (+)RNA plant viruses, e.g., potyviruses. Conserved domains, which have been identified in the 5′ UTR of potyviruses [Maiss et al., 1989; Turpen, 1989] are absent from both BaMMV RNA2 [this work] and BaYMV RNA2 [Davidson et al., 1991].

Recently, Kashiwazaki et al. [1991] reported about two short stretches of 7 nts (AAAGCAA and UUGCUUU) identical in the 5′ region of BaYMV RNA1 and RNA2 that have the potential to base-pair. While some short sequences of potential base-pairing can also be identified by computer analysis in the 5′ UTR of BaMMV RNA2, they do not resemble the sequences described for BaYMV.

BaMMV RNA2 contains a single large ORF between nts 141 and 2819. Four AUG start codons are present at the 5′ terminus of the ORF, but only the first, starting at nt 141 is in good agreement with the consensus sequence for initiation of translation of eukaryotic mRNAs [Kozak, 1986; Staden, 1984]. The putative polyprotein encoded by this ORF comprises 893 amino acids and has a predicted molecular weight of 98 kDa. It shows only low similarity to the amino acid sequence of BaYMV RNA2. The most prominent homology is found in the N terminus of the polyprotein (Fig. 2). The amino acid motif GYCY, which occurs in the polyprotein encoded by the BaYMV RNA2, and in the HC-Pro of aphid-transmissible potyviruses, was identified to contain a proteolytically active cysteine [Oh and Carrington, 1989]. In BaMMV this domain has altered to GFCY and the putative active cysteine residue occurs at position 117. Furthermore, a histidine residue known as being essential for cleavage activity of HC-Pro of potyviruses has also been found in the polyprotein of BaMMV RNA2, at amino acid position 189. As in BaYMV, it is located 72 amino acids downstream of the cysteine residue of the GFCY motif. The surrounding amino acids of this histidine residue are nearly identical in the corresponding protein of the Streatley isolate of BaMMV [Andersen et al., 1993].

```
5' AAAAUAAAACUUCACAAAAACCCAACGACUUAAAACGAACACAAUUAAGCAAACAAACAA 3'
          10        20        30        40        50        60

5' ACACCAAACGCACACAACGAAUUACAGCACACAGCUGUAAAGCGCGAACUCAUUCACCGC 3'
          70        80        90       100       110       120

                        M  M  M  N  S  M  I  R  Q  G  W  Q  Q   13
5' AAUCUUUGCAUUGAGCCAAGAUGAUGAUGAACUCCAUGAUCCGCCAAGGAUGGCAGCAGG 3'
         130       140       150       160       170       180

   V  L  R  R  F  S  I  P  T  S  G  D  R  L  I  V  S  N  S  T   33
5' UGCUGAGGCGCUUUUCCAUCCCAACAUCCGGAGAUAGACUCAUCGUUUCCAACUCCACCG 3'
         190       200       210       220       230       240

   D  Q  P  I  G  L  F  G  A  F  D  T  S  L  Q  T  L  S  Q  V   53
5' ACCAACCCAUCGGCCUUUUCGGAGCGUUCGACACAUCGCUGCAAACACUCUCACAAGUCA 3'
         250       260       270       280       290       300

   T  N  D  P  E  V  L  K  Q  K  S  N  I  P  T  H  L  D  V  A   73
5' CCAAUGAUCCUGAGGUCUUGAAGCAGAAGUCCAACAUCCCCACGCACCUGGACGUGGCCU 3'
         310       320       330       340       350       360

   S  V  L  E  A  S  P  R  S  F  P  W  V  F  L  T  N  S  F  C   93
5' CCGUUCUUGAAGCAAGCCCACGGUCCUUCCCAUGGGUUUUUCUCACGAACUCAUUCUGCA 3'
         370       380       390       400       410       420

   T  F  G  G  S  I  H  A  Q  N  L  Q  A  F  A  T  A  E  F  K  113
5' CCUUCGGUGGUUCGAUACAUGCUCAGAAUCUUCAAGCCUUCGCAACUGCCGAGUUUAAGA 3'
         430       440       450       460       470       480

   S  G  F  C  Y  M  N  L  L  I  P  L  S  F  D  I  I  D  A  H  133
5' GCGGGUUCUGUUACAUGAACCUGCUAAUACCACUCAGCUUUGAUAUCAUCGACGCACACG 3'
         490       500       510       520       530       540

   A  D  S  F  R  G  F  V  E  Q  L  P  D  T  L  G  A  Y  P  S   153
5' CAGAUUCUUUUCGCGGCUUUGUUGAACAACUUCCUGACACGUUAGGCGCCUACCCUUCAC 3'
         550       560       570       580       590       600

   L  S  M  V  L  N  V  M  L  H  A  A  T  R  F  P  E  I  V  A   173
5' UCAGCAUGGUGCUAAACGUCAUGUUGCACGCAGCCACGCGUUUCCCCGAGAUCGUCGCCU 3'
         610       620       630       640       650       660

   S  P  I  P  T  I  A  F  D  A  E  S  L  Q  F  H  V  T  D  K   193
5' CCCCUAUCCCAACCAUAGCAUUUGACGCCGAAUCACUGCAGUUCCAUGUUACUGAUAAAC 3'
         670       680       690       700       710       720

   R  G  V  P  G  M  W  N  I  L  K  A  C  R  V  Y  E  L  L  S   213
5' GGGGUGUCCCUGGCAUGUGGAACAUACUCAAAGCCUGCCGUGUGUACGAGCUUCUUUCGC 3'
         730       740       750       760       770       780

   L  A  A  D  G  I  G  C  E  Y  M  L  Y  P  V  G  A  A  P  Q   233
5' UUGCAGCAGAUGGUAUAGGUUGCGAGUACAUGCUGUACCCUGUUGGUGCUGCACCCCAAU 3'
         790       800       810       820       830       840

   Y  S  F  W  K  K  S  M  D  H  F  T  S  D  R  F  V  E  F  L   253
5' ACUCCUUCUGGAAGAAAUCCAUGGACCACUUCACAUCCGACCGGUUCGUGGAAUUCCUAG 3'
         850       860       870       880       890       900

   A  M  Q  D  L  L  A  S  A  L  E  Q  D  Y  A  T  H  D  A  R   273
5' CCAUGCAGGAUCUGCUCGCCAGUGCACUUGAACAAGAUUAUGCAACACAUGACGCGCGCG 3'
         910       920       930       940       950       960

   D  A  L  L  S  A  L  Q  N  A  G  Y  T  N  V  V  A  R  E  R   293
5' AUGCACUACUUUCAGCUCUCCAGAACGCAGGUUACACAAACGUUGUUGCAAGGGAGAGGA 3'
         970       980       990      1000      1010      1020

   R  F  P  N  G  H  D  P  S  I  V  W  L  N  L  S  E  A  P  I   313
5' GGUUUCCAAAUGGCCACGACCCUAGCAUUGUCUGGUUGAACCUGAGCGAAGCCCCCAUUU 3'
        1030      1040      1050      1060      1070      1080

   S  E  K  L  T  E  L  K  R  Y  L  L  V  G  H  R  S  D  D  T   333
5' CAGAGAAGCUAACUGAACUUAAACGGUACCUCCUAGUCGGACAUAGGAGCGACGACACUG 3'
        1090      1100      1110      1120      1130      1140
```

Fig. 1. Nucleotide sequence of BaMMV RNA2 and the amino acid sequence derived from the single large ORF starting at nucleotide 141.

```
   A  D  I  T  H  N  V  H  Q  H  V  F  E  V  L  K  T  M  S  V   353
5' CUGAUAUCACACACAACGUGCACCAGCACGUUUUUGAAGUCCUCAAGACCAUGUCUGUGC 3'
        1150       1160       1170       1180       1190       1200

   Q  F  S  K  T  T  N  A  Y  N  R  A  R  F  E  V  N  H  K  V   373
5' AGUUUAGCAAGACAACGAACGCCUACAACCGUGCCCGCUUUGAAGUCAACCACAAAGUCA 3'
        1210       1220       1230       1240       1250       1260

   I  W  N  A  E  Y  G  R  G  P  Q  Q  N  A  E  L  E  A  L  V   393
5' UAUGGAACGCAGAAUACGGACGCGGCCCUCAACAGAACGCCGAGCUCGAGGCUCUUGUAC 3'
        1270       1280       1290       1300       1310       1320

   L  F  L  N  R  Q  S  L  E  I  E  N  I  L  H  R  T  T  S  P   413
5' UCUUUCUUAAUAGGCAAUCGCUUGAAAUUGAGAACAUCUUGCACAGAACAACUAGUCCUG 3'
        1330       1340       1350       1360       1370       1380

   V  V  V  T  S  W  K  P  D  V  P  P  A  A  P  E  I  K  E  E   433
5' UUGUUGUCACCAGCUGGAAACCCGAUGUUCCCCCUGCUGCACCUGAAAUCAAAGAGGAGG 3'
        1390       1400       1410       1420       1430       1440

   E  P  T  H  A  I  A  T  P  I  T  E  A  P  S  H  V  T  P  V   453
5' AACCCACGCACGCAAUCGCAACACCAAUAACUGAGGCGCCAUCACAUGUCACUCCUGUUG 3'
        1450       1460       1470       1480       1490       1500

   E  V  V  N  L  P  P  T  R  S  Y  W  A  E  T  L  V  G  I  L   473
5' AAGUUGUGAACCUGCCACCAACGCGCUCCUAUUGGGCUGAGACACUUGUCGGUAUCCUCA 3'
        1510       1520       1530       1540       1550       1560

   T  A  I  L  G  T  V  F  A  F  L  T  R  A  L  I  R  A  K  R   493
5' CUGCAAUCCUGGGAACGGUAUUCGCCUUCCUCACUCGAGCACUCAUUCGUGCAAAGAGAU 3'
        1570       1580       1590       1600       1610       1620

   L  R  R  K  S  T  F  P  W  V  T  L  N  S  G  D  D  D  D  D   513
5' UGCGGAGGAAAUCCACCUUCCCCUGGGUGACGCUAAACUCUGGAGAUGACGAUGAUGACC 3'
        1630       1640       1650       1660       1670       1680

   Q  S  G  G  G  G  G  G  P  Q  T  P  G  G  Q  P  P  V  P  H   533
5' AAUCUGGGGGGGGAGGAGGUGGCCCGCAAACACCUGGCGGGCAGCCCCCAGUCCCGCACA 3'
        1690       1700       1710       1720       1730       1740

   T  R  G  T  H  Q  S  R  F  S  V  Q  D  I  A  S  D  T  S  L   553
5' CGCGUGGAACGCACCAGUCGCGCUUCUCCGUGCAGGACAUAGCAUCCGACACCAGUCUGC 3'
        1750       1760       1770       1780       1790       1800

   L  S  V  D  L  D  E  D  T  L  S  Q  Y  D  E  T  F  Q  K  I   573
5' UAAGUGUCGAUCUCGACGAGGACACGCUCUCCCAGUAUGAUGAAACCUUCCAAAAGAUCC 3'
        1810       1820       1830       1840       1850       1860

   R  R  A  L  F  E  T  S  F  A  D  I  L  Q  N  S  A  R  W  I   593
5' GCCGUGCGCUCUUCGAAACCAGUUUUGCAGACAUCCUGCAAAACUCUGCUCGUUGGAUCU 3'
        1870       1880       1890       1900       1910       1920

   S  T  L  E  A  M  A  L  A  D  G  N  A  P  Y  T  L  L  A  Q   613
5' CCACGCUGGAAGCGAUGGCUCUCGCAGAUGGCAAUGCUCCUUACACACUUCUUGCGCAGU 3'
        1930       1940       1950       1960       1970       1980

   Y  L  N  G  I  E  E  A  Y  T  N  F  R  N  T  G  H  I  S  R   633
5' AUCUCAACGGGAUUGAGGAGGCAUACACGAACUUUCGCAACACAGGUCACAUCUCCCGCG 3'
        1990       2000       2010       2020       2030       2040

   A  T  L  S  G  F  F  A  L  E  D  N  L  R  A  A  G  I  A  F   653
5' CAACACUUUCAGGCUUCUUCGCCUUGGAAGAUAAUCUGCGUGCAGCUGGCAUAGCUUUUG 3'
        2050       2060       2070       2080       2090       2100

   G  T  T  T  P  T  Q  T  I  Q  N  Q  F  A  D  S  P  A  R  R   673
5' GGACCACAACACCCACGCAAACCAUACAGAACCAAUUUGCCGACUCCCCAGCCCGUCGAU 3'
        2110       2120       2130       2140       2150       2160

   W  K  T  R  F  E  Q  I  A  C  E  L  G  D  A  S  I  K  S  L   693
5' GGAAAACACGAUUUGAACAGAUUGCAUGUGAACUGGGCGAUGCUAGCAUCAAAUCGCUUG 3'
        2170       2180       2190       2200       2210       2220

   A  D  L  A  D  I  I  D  T  E  R  E  R  G  D  L  T  Q  F  D   713
5' CAGAUCUUGCAGACAUCAUCGACACUGAGCGGGAGAGGGGCGACUUGACUCAGUUUGAUG 3'
        2230       2240       2250       2260       2270       2280
```

Fig. 1. Continued.

```
   V  L  A  A  S  S  I  S  S  L  C  R  A  V  R  I  I  S  D  T   733
5' UUCUCGCAGCCUCUUCCAUCAGCAGCCUCUGUCGUGCUGUGCGCAUCAUCAGUGACACCA 3'
        2290       2300       2310       2320       2330       2340

   T  D  P  N  A  Q  L  A  L  V  E  N  A  T  A  M  Q  N  N  I   753
5' CCGAUCCAAAUGCUCAAUUAGCCCUUGUUGAAAACGCAACGGCUAUGCAGAACAACAUCA 3'
        2350       2360       2370       2380       2390       2400

   N  A  I  L  G  T  N  V  S  I  P  F  L  S  A  T  R  R  L  L   773
5' ACGCCAUACUGGGCACGAACGUUAGCAUACCUUUCCUCAGUGCGACGCGUCGUUUACUCA 3'
        2410       2420       2430       2440       2450       2460

   I  T  R  R  V  Q  E  A  G  A  E  S  R  S  G  A  T  P  E  T   793
5' UCACACGCCGUGUUCAAGAAGCCGGGGCGGAGAGCCGCUCUGGAGCAACGCCUGAAACCG 3'
        2470       2480       2490       2500       2510       2520

   V  Q  Q  L  A  D  A  E  L  A  K  I  V  S  E  A  N  M  Y  N   813
5' UUCAGCAACUGGCUGACGCAGAGCUGGCAAAAAUUGUGAGCGAGGCAAACAUGUACAAUG 3'
        2530       2540       2550       2560       2570       2580

   E  M  A  A  S  Q  R  D  I  A  N  A  T  R  E  A  T  I  R  E   833
5' AGAUGGCCGCCAGCCAAAGGGACAUCGCCAACGCAACACGCGAAGCCACCAUCCGCGAGC 3'
        2590       2600       2610       2620       2630       2640

   H  V  L  S  P  V  N  A  L  A  N  V  G  M  A  A  A  F  F  R   853
5' ACGUGCUCAGCCCCGUCAACGCUUUGGCAAAUGUUGGCAUGGCUGCUGCGUUCUUCCGAU 3'
        2650       2660       2670       2680       2690       2700

   S  G  G  L  R  S  R  A  F  N  P  A  M  P  T  M  P  G  G  P   873
5' CCGGUGGUUUGCGCUCAAGGGCCUUCAACCCCGCAAUGCCAACAAUGCCCGGUGGUCCUG 3'
        2710       2720       2730       2740       2750       2760

   A  A  A  G  R  P  M  F  Q  A  F  R  G  R  G  H  R  L  N  R   893
5' CUGCUGCGGGGCGGCCAAUGUUCCAAGCCUUCAGGGGACGAGGGCAUCGUCUUAAUAGGU 3'
        2770       2780       2790       2800       2810       2820

   *
5' AGGCAUCAGAACUCUGUGCUCUGGCACAGCACUGAACGAGCGACACGCGCUCGUUGUGAA 3'
        2830       2840       2850       2860       2870       2880

5' GAUCCAAUUCCACGCCCAGCGUGCCACUGAACAAACGAAACACGUUUGUUGUGAUAAUUA 3'
        2890       2900       2910       2920       2930       2940

5' AUUAAUUUUUGCGCAGACUUUGUCUGUGUUUUCUAGAGUGUGCUGCUCUGCAGCUGCAUU 3'
        2950       2960       2970       2980       2990       3000

5' GGAGGGUUUCACGGGAACGACAUCAUCACACCCCAUCCGCAAACCGCCUGCAUUCACAUC 3'
        3010       3020       3030       3040       3050       3060

5' UUUGUGCCACUUUAAAUACUCAUACGCAUAGGCAAGUGCCUUUCUGGCAGAGCAUCCCCG 3'
        3070       3080       3090       3100       3110       3120

5' AGCAUCGACAUUCACGCGAUUGCAAUCAACUGGCUUAAUCCCCGGUUUUGCAAUCAAUAG 3'
        3130       3140       3150       3160       3170       3180

5' UGAGUUCUACCUCCGGACGUGUUCCCCUGGAACUGCAAUCCUCGUUAUGUUGUCGCUGAU 3'
        3190       3200       3210       3220       3230       3240

5' UACUAUCAGCUCGUAAACCAUGAUAUCCACCCACAAUCAGCUACUCGUGUACUAACAGUA 3'
        3250       3260       3270       3280       3290       3300

5' CCGGACACGCCUAUCCUGCACUUAAGCCUGCAGGACGCUACCAGGUCCGCCUACUGUUAC 3'
        3310       3320       3330       3340       3350       3360

5' CAUUUGUGUAGUUUGUUGUGGCUCCACCUCAUAUCUAAAUUCAGUACCUGCAUACAAAAC 3'
        3370       3380       3390       3400       3410       3420

5' GUCAUGCAUCCCUGUUCAACAUGAGCACGGUUUGGGAAGAAGCAAGUCAUCAUAGACUGC 3'
        3430       3440       3450       3460       3470       3480

5' UGGCAUUGUUAGGUUGACAAUGUUUAUGUACACAGGAACGUGAC 3'
        3490       3500       3510       3520
```

Fig. 1. Continued.

```
BaMMV 112  F K S G F C Y M N L L I P L S F D I I D
BaYMV 138  * A H * Y * * L S * F * * * * * R * T P

           A H A D S F R G F V E Q L P D T L G A Y
           E N * R * * S R * L * * * * * I * * * *

           P S L S M V L N V M L H A A T R F P E I
           * T * A S L Y K T * * F * V R L * * * V

           V A S P I P T I A F D A E S L Q F H V T
           L Q A * * * I * * K R P G V * * * * * S

           D K R G V P G M W N I L K A C R V Y E L
           * A * * L * P S * F P M * C G S * A S F

           L S L A A D G I G C E Y M L Y P V G A 230
           I A * I T N N L N S D L L N G I * * S 256
```

Fig. 2. Comparison of amino acids 112 to 230 of BaMMV RNA2 with the corresponding region of the polyprotein of the German isolate of BaYMV [Davidson et al., 1991]; amino acid residues identical for BaMMV and BaYMV are marked by asterisks (*) in the BaYMV sequence; boxed sequences represent the proteolytically active regions of the putative proteinase and the putative cleavage site of the polyprotein.

The putative G/S cleavage site in the polyprotein of BaYMV RNA2 is replaced by a G/A dipeptide in BaMMV at positions 229/230. In addition, as for potyviruses and BaYMV the G is preceded by a valine.

Cleavage at this site will lead to a 25 kDa protein, a putative proteinase (28 kDa in BaYMV) and a 73 kDa protein of unknown function (70 kDa in BaYMV). No other ORF, starting with an AUG and a coding capacity for a peptide larger than 12 kDa, could be detected.

The 3′ non-coding region of BaMMV RNA2 is 705 nts long and, as shown in Fig. 3, there is a good homology between the German isolate and the published 141 nts from the extreme 3′ terminus of RNA2 of the Streatley isolate of BaMMV [Andersen et al., 1993]. The 3′ non-coding region contains the sequence UAUGU at positions 3505 to 3509. This sequence was originally described at the DNA level to be a polyadenylation signal in yeast and necessary for transcription termination [Zaret and Sherman, 1982]. It was first detected in the sequence of a plant virus RNA by Maiss et al. [1989] and similar putative signals have been identified in a number of other plant virus RNAs. However, as yet the proposed function

```
a) 3385 5' CACCUCAUAUCUAAAUUCAGUA-CCUGCAUACAAAACGUCAUGCAUCCCUGUUCAACAUGA
b) -141 5' *U***********-*******A**********-**********A**********

a)         GCACGG-UUUGGGAAGAAGCAAGUCAUCAUAGACUGCUGGCAUUGUUAGGUUGACAAUGUU
b)         ******G*****************************************************C*

a)         UAUGUACACAGGAACG-UGAC 3' 3524
b)         **********A**** 3' 0
```

Fig. 3. Comparison of the nucleotide sequence of the 3′-terminal parts of RNA2 of the BaMMV isolates from (a) Aschersleben, Germany and (b) Streatley, U.K. [Andersen et al., 1993].

of this signal has not been confirmed experimentally. Infectious full-length clones should be a good tool to elucidate the biological function of this sequence.

The nucleotide sequence data reported here will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X75933.

Acknowledgements

We should like to thank Mrs Doris Walther for excellent technical support and Drs H. Huttinga, F. van der Wilk and G. Proeseler for their engagement at the initial stage of this work. UT was granted by the Deutsche Forschungsgemeinschaft (Pr 420/1-1).

References

Adams MJ, Batista MF, Swaby AG and Antoniw JF (1989) Fungally-transmitted viruses of cereals in the UK. Eppo Bull. 19: 573–577

Andersen JF, Davies JW and Coutts RHA (1993) Further evidence for the inclusion of barley mild mosaic virus (BaMMV) in the baymovirus group. J. Phytopathology 139: 48–56

Davidson AD, Pröls M, Schell J and Steinbiss H-H (1991) The nucleotide sequence of RNA2 of barley yellow mosaic virus. J. gen. Virol. 72: 989–993

Ehlers U and Paul H-L (1986) Characterization of the coat proteins of different types of Barley yellow mosaic virus by polyacrylamide gel electrophoresis and electro-blot immunoassay. J. Phytopathology 115: 294–304

Gubler U and Hoffman BJ (1983) A simple and efficient method for generating cDNA libraries. Gene 25: 263–269

Huth W (1986) Isolierung mehrerer Stämme des Gelbmosaikvirus der Gerste (barley yellow mosaic virus, BaYMV). Mitt. Biol. Bundesanstalt Berlin 232: 384

Huth W and Adams MJ (1990) Barley yellow mosaic virus (BaYMV) and BaYMV-M, two different viruses. Intervirology 31: 38–42

Huth W, Lesemann D-E and Paul H-L (1984) Barley yellow mosaic virus: purification, electron microscopy, serology and other properties of two types of the virus. Phytopath. Z. 111: 37–54

Kashiwazaki S, Hayano Y, Minobe Y, Omura T, Hibino H and Tsuchizaki T (1989) Nucleotide sequence of the capsid protein gene of barley yellow mosaic virus. J. gen. Virol. 70: 3015–3023

Kashiwazaki S, Minobe Y and Hibino H (1991) Nucleotide sequence of barley yellow mosaic virus RNA2. J. gen. Virol. 72: 995–999

Kashiwazaki S, Minobe Y, Omura T and Hibino H (1990) Nucleotide sequence of barley yellow mosaic virus RNA1: a close evolutionary relationship with potyviruses. J. gen. Virol. 71: 2781–2790

Kashiwazaki S, Nomura K, Kuroda I and Hibino H (1992) Sequence analysis of the 3′-terminal halves of RNA1 of two strains of barley mild mosaic virus. J. gen. Virol. 73: 2173–2181

Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon of eucaryotic ribosomes. Cell 44: 283–292

Maiss E, Timpe U, Brisske A, Jelkmann W, Casper R, Himmler G, Mattanovich D and

Katinger HWD (1989) The complete nucleotide sequence of plum pox virus RNA. J. gen Virol. 70: 513–524

Oh C-S and Carrington JC (1989) Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. Virology 173: 692–699

Peerenboom E, Pröls M, Schell J, Steinbiß H-H and Davidson AD (1992) The complete nucleotide sequence of RNA1 of a German isolate of barley yellow mosaic virus and its comparison with a Japanese isolate. J. gen. Virol. 73: 1303–1308

Sanger F, Nicklen S and Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl. Acad. Sci. USA 74: 5463–5467

Schlichter U, Sohn A, Peerenboom E, Schell J and Steinbiß H-H (1993) Molecular analysis of the capsid protein gene of a german isolate of barley mild mosaic virus. Plant Cell Rep. 12: 237–240

Staden R (1984) Computer methods to locate signals in nucleic acid sequences. Nucl. Acids Res. 12: 505–519

Turpen T (1989) Molecular cloning of a potato virus Y genome: nucleotide sequence homology in non-coding regions of potyviruses. J. gen Virol. 70: 1951–1960

Usugi T, Kashiwazaki S, Omura T and Tsuchizaki T (1989) Some properties of nucleic acids and coat proteins of soil-borne filamentous viruses. Ann. Phytopathol. Soc. Jpn. 55: 26–31

Zaret KS and Sherman F (1982) DNA sequence required for efficient transcription termination in yeast. Cell 28: 563–573