The tymobox, a sequence shared by most tymoviruses: its use in molecular studies of tymoviruses

Shouwei Ding, Jennifer Howe[1], Paul Keese[1], Anne Mackenzie, Drew Meek[2], Marlene Osorio-Keese[3], Mary Skotnicki, Pattana Srifah, Marjo Torronen and Adrian Gibbs
Research School of Biological Sciences, The Australian National University, Canberra, ACT 2601, [1]CSIRO Division of Plant Industry, Black Mountain, Canberra, ACT, [2]National Biological Standards Laboratory, Canberra, ACT and [3]Patents Trademarks & Designs Office, Scarborough House, Phillip, Canberra, ACT, Australia

Received November 23, 1989; Revised and Accepted January 22, 1990

ABSTRACT

The 5′-terminal sequences of the virion protein mRNAs of ononis yellow mosaic and kennedya yellow mosaic tymoviruses were determined, and also the positions in the genomes of the transcription initiation sites of those mRNAs. Comparisons of the available genomic sequences of tymoviruses revealed two conserved regions, one at the initiation site and another longer sequence of sixteen nucleotides to the 5′ side of it. The longer sequence, which we call the tymobox, was tested as a target for a designed ribozyme, which cleaved appropriate genomic fragments of three tymoviruses. A synthetic oligonucleotide with sequence complementary to the tymobox was shown to be a tymovirus-specific probe for diagnosing and identifying tymoviruses, except for wild cucumber mosaic tymovirus. The tymobox sequence was also used as a primer for the second strand DNA synthesis of dsDNA representing the virion protein gene of cacao yellow mosaic tymovirus, a tymovirus with unknown sequence. Thus, the tymobox is a useful tool in molecular studies of tymoviruses.

INTRODUCTION

Tymoviruses are a group of plant viruses which have single-stranded messenger sense genomic RNAs about 6 kb in length (1). The particles of tymoviruses are isometric and of two types with different sedimentation coefficients. The particles of the slow, or top, component are not infectious. Those of TYMV are empty shells whereas those of other tymoviruses contain subgenomic RNA that is the messenger for the virion protein. The particles of the bottom component contain both genomic and subgenomic RNAs and are infectious (1). The complete sequences of the genomic RNAs of TYMV (2, 3, and D. Meek, unpublished data), EMV-Trin (4), OYMV-Tin (5), KYMV-JB (6) and ELV (Srifah *et al.*, in preparation) have been recently determined. Partial genomic sequences of several other tymoviruses, including DMV and BdMV-Eur (J. Howe, unpublished data), APLV-Hu (M. Osorio-Keese, unpublished data), TYMV-Roth (M. Torronen, unpublished data), and CYVV, WCuMV, KYMV-PD and KYMV-BP (A. Mackenzie, unpublished data) have also be determined in our laboratory. It was shown that all of these genomes contain three open-reading frames (ORFs), two ofwhich have their initiation codons close to the 5′-end of the genome and overlap, whereas the virion protein gene is located at the 3′ end portion of the genome and is expressed via a subgenomic mRNA (7). Sequence analysis has shown that tymoviral encoded replicase proteins (RP) are more similar to those encoded by members of the supergroup of Sindbis-like viruses than those of picorna-like viruses (3). In particular, tymoviral RPs were found to be closely related to those of potexviruses and carlaviruses (6).

In this paper, we report the 5′ terminal sequences of the subgenomic RNAs of OYMV-Tin and KYMV-JB as well as their initiation sites in the corresponding genomic RNA. Comparisons of the tymovirus genomic sequences determined so far have revealed a conserved sequence of 16 nucleotides, which we call the tymobox, located a few nucleotides to the 5′ side of the start of the subgenomic RNA messenger. The tymobox, which may be an important element of the subgenomic RNA promoter, has been shown experimentally to be a specific target for, firstly, cleavage by a designed ribozyme, also for a specific hybridization probe for diagnosis and identification of most tymoviruses, and for a primer for cloning the virion protein gene of most tymoviruses.

MATERIALS AND METHODS

Viruses and viral RNAs

The sources of all tymoviruses used in this study are the same as in (8) and viral particles and RNAs were purified as previously described (9).

Oligonucleotides

Eight oligonucleotides: 5′-GACAGATGAGTCCCA-3′ (P1), 5′-GGACTTGGTGTTCAC-3′ (P2), 5′-GAAGCAATTCA$^{G}_{A}$ACTC-3′ (P3), 5′CCCTCGAGT$^{C}_{A}$TGAATTGCTTC-3′ (P4), 5′-AATTCTAGAGTTTCGTCCTTTAGGGACTCGTCAGT-GAATTGCTTCCTGCA-3′ (P5), 5′-GGAAGCAATTCAC-TGACGAGTCCCTAAAGGACGAAACTCTAG-3′ (P6), 5′-GGGGTTCCCGTTGACCCACG-3′ (P7), and 5′-CTTTAA-

CGTCGCGAAGAC-3′ (P8) were kindly synthesized by M. Gibbs (P1, P2 and P6) in the Division of Plant Industry, CSIRO and G. Mayo (P3, P4 and P5) in the John Curtin School of Medical Research, ANU. The sequences of P1 and P2 are complementary to nucleotides 5536–5551 of the OYMV-Tin genome (5) and nucleotides 5747–5761 of the KYMV-JB genomes (6) respectively. The sequence of P3 is complementary to that of the tymobox, and that of P4 is the same as the tymobox except a XhoI site was incorporated. P5 and P6 are complementary and overlapping oligonucleotides and contain the catalytic domain from subterranean clover mottle virus satellite RNA (J. Haseloff, personal communication). P7 is complementary to nucleotides 6313–6330 (excluding the first two nucleotides at the 5′ end of the P7) of the EMV-Trin genomic RNA (4), and P8 is complementary to a sequence at the 3′ terminal region of BdMV-Eur genome obtained by direct RNA enzymatic sequencing (J. Howe, unpublished result).

Determining the 5′ terminal sequences of subgenomic RNAs

The two synthetic primers, P1 and P2, were used for direct sequencing of both genomic and subgenomic RNAs of OYMV-Tin and KYMV-JB by the dideoxynucleotide chain termination method using reverse transcriptase (10).

Cloning of the ribozyme

About 1 μg of each of the two oligonucleotides, P5 and P6, was phosphorylated (11) with ATP and T4 polynucleotide kinase (New England Biolabs). The phosphorylated oligonucleotides were annealed in 10 μl sterile H_2O by heating at 80°C for 2 minutes and cooling slowly to room temperature. A plasmid, pTyRz, was generated by ligating 0.5 μl of the annealed double-stranded oligonucleotides (approximately 20 ng) into the pTZ19 vector that had been hydrolysed by endonucleases EcoRI and PstI, and dephosphorylated by alkaline phosphatase (Boehringer-Mannheim). Thus, the ribozyme-derived DNA sequence was located to the 3′ side of the promoter for T7 RNA polymerase in pTyRz. The ligation mixture was used to transform competent JM101 cells of *E. coli* (11).

Cloning of the substrate sequences

cDNAs to the parts of the genomic sequences of OYMV-Tin, KYMV-JB and BdMV-Eur, that contain the tymobox, were cloned and RNA transcripts from these clones were used as substrates for the cleavage reaction. These parts of the genomes, genomic positions of which are illustrated in Fig. 1, were selected in the following ways:

(i) *OYMV-Tin*. Synthetic random oligonucleotide hexamers (Bresatec, Australia) were used to prime the synthesis of cDNA to OYMV-Tin genomic RNA (12). A DNA fragment of 1181 bps (representing nucleotides 4834–6016 of the genome) was obtained by hydrolyzing the synthesized dsDNA with SphI and HindIII, and was selected by electrophoresis in 5% polyacrylamide gel (11). This fragment was then cloned between the sites of SphI and HindIII of pTZ18 (11) and called pOY4. Using one PstI site in the multiple cloning region of the vector and the other in the insert (at nucleotide 5217 of the genome), the smaller PstI fragment (corresponding to nucleotides 4834–5217 in the genome) was removed from pOY4 by hydrolyzing with endonuclease PstI, and the larger fragment was circularized. This was named pOY41 and represented nucleotides 5218–6014 of the OYMV genome and hence contained the tymobox (5).

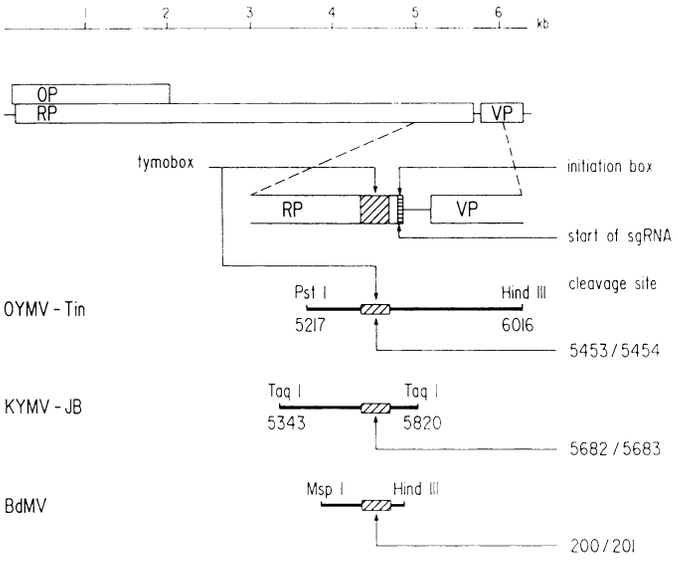

Fig. 1 Diagram illustrating the positions of the three tymoviral genomic RNA fragments targeted by the tymoribozyme. Top line represents the tymovirus genome (its scale in kilobases) and three possible encoded proteins (open boxes) are shown below. The junction region of RP (replicase protein) and VP (virion protein) is enlarged to show the location of the tymobox. OP—overlapping, out-of-phase protein; sgRNA—subgenomic RNA.

(ii) *KYMV-JB*. A M13mp18 single-stranded DNA clone, named JB249, was constructed during the course of sequence determination of the KYMV-JB genome (6). It contains a TaqI fragment of the viral cDNA which is 477 nucleotides long and in the genomic sense. This fragment represents nucleotides 5343–5820 of KYMV-JB genomic RNA and contains a tymobox with one difference (C⇒U in the fifth position) from the commonest tymobox sequence. The 17-mer M13 universal primer was used to initiate second-strand DNA synthesis of JB249 using 1 unit of *E. coli* DNA polymerase I (large fragment, Bresatec, Australia) in the presence of 0.5 mM dNTP's, 10μCi α-^{32}P-dCTP (Amersham), 10 mM Tris-HCl (pH 8.0), and 10 mM $MgCl_2$. The reaction mixture was incubated at 37°C for 30 minutes before the reaction was stopped by shaking with phenol/chloroform. The DNA was precipitated by ethanol and NaoAC, and hydrolysed with endonucleases HindIII and EcoRI. The smaller EcoRI-HindIII fragment (less than 500 bps) was separated from the rest by electrophoresis in 5% polyacrylamide gel, and eluted from the gel slice by soaking overnight in the buffer containing 10 mM Tris-HCl, pH8.0, 1 mM EDTA and 1% SDS. The purified DNA fragment was ligated into the pTZ18 vector which had been hydrolysed by EcoRI and HindIII and dephosphorylated by calf intestinal phosphorylase. The transformation was done in the way described above for the OYMV-Tin fragment. The resulting clone was called pJB249.

(iii) *BdMV-Eur*. An M13mp18 clone of BdMV-Eur genome named BdMV162 was kindly provided by J. Howe. This single-stranded DNA clone contained a MspI-HindIII viral cDNA fragment which is of complementary orientation and approximately 290 bps in length including the tymobox. As with the KYMV-JB fragment, the viral-derived sequence in BdMV162 was subcloned between the HindIII and EcoRI sites of the plasmid pTZ19 to produce pBM162.

In vitro transcription and cleavage

The cloned DNA fragments were transcribed in vitro as described by Melton et al. (13). The DNA of pOY41 and pJB249 were first linearised by hydrolyzing with HindIII, whereas the DNA of pTyRz and pBM162 were linearised with EcoRI. The plus strand RNAs were transcribed from the DNAs using 500 units ml^{-1} T7 RNA polymerase (Promega) in 40 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 1000 units ml^{-1} RNasin (Promega), 500 μM ATP, CTP, GTP, 24 μM UTP with 20 μCi α-^{32}P-UTP (Amersham). The RNA transcripts were separated from their DNA templates by electrophoresis in 8% polyacrylamide gel containing 7 M urea. Bands containing the RNA transcripts were cut from the gels and the RNA eluted by soaking (11). RNA transcripts were all resuspended in sterile water before treating with the ribozyme. The most efficient conditions known for *in vitro* cleavage (14) were used. The substrate and ribozyme RNAs were incubated in 50 mM Tris-HCl pH 8.0, 20 mM $MgCl_2$ at 50°C for 60 minutes. The products were then fractionated in an 8% polyacrylamide gel containing 7 M urea and were autoradiographed.

Hybridization procedure

The oligonucleotides were radioactively labelled with polynucleotide kinase and γ-^{32}P-ATP and purified by electrophoresis in a polyacrylamide gel. Purified viral RNAs (1 μg) were blotted onto prewetted Zeta-Probe blotting membrane and hybridized with the probe at 43°C as described (15). The membrane was washed with 2×SSC/0.1% SDS, 0.5×SSC/0.1% SDS and 0.1×SSC/0.1% SDS, respectively at room temperature for 30 minutes, and then autoradiographed.

Cloning the tymoviral virion protein genes

CoYMV genomic RNA was polyadenylated with ATP and poly (A) polymerase (Bresatec). A kinased synthetic primer, dT_8G, was used to prime the first cDNA synthesis on the polyadenylated CoYMV RNA, whereas the specific primers P7 and P8 were used to initiate the first cDNA synthesis on the genomic RNAs of EMV and BdMV-Eur, respectively (12). The virion protein genes were amplified (16) by polymerase chain reaction (PCR) using as primers dT_8G and P4 on the cDNA transcribed from CoYMV, P7 and P4 on that from EMV-Trin, and P8 and P4 on that from BdMV-Eur. The resulting dsDNAs were fractionated from a 5% polyacrylamide gel and cloned into SmaI site of M13mp18 and sequenced by the dideoxynucleotide chain termination method (17).

RESULTS

5′ terminal sequences of the subgenomic RNAs of OYMV-Tin and KYMV-JB

Parts of the subgenomic and genomic RNAs of OYMV-Tin and KYMV-JB were sequenced using the chosen primers and compared directly in the same sequencing gels. Although the 5′ terminal nucleotide of the subgenomic RNA of OYMV-Tin was not clear in the sequencing gel of the subgenomic RNA (Fig. 2), its position was obvious in the genomic RNA tracks as there was a change in band pattern, probably caused by the presence of a small amount of subgenomic RNA in the genomic RNA preparation. The presence of bands in all five lanes, especially in lane N where no dideoxynucleotide was added to the sequencing reaction, indicates that these cDNA transcripts are

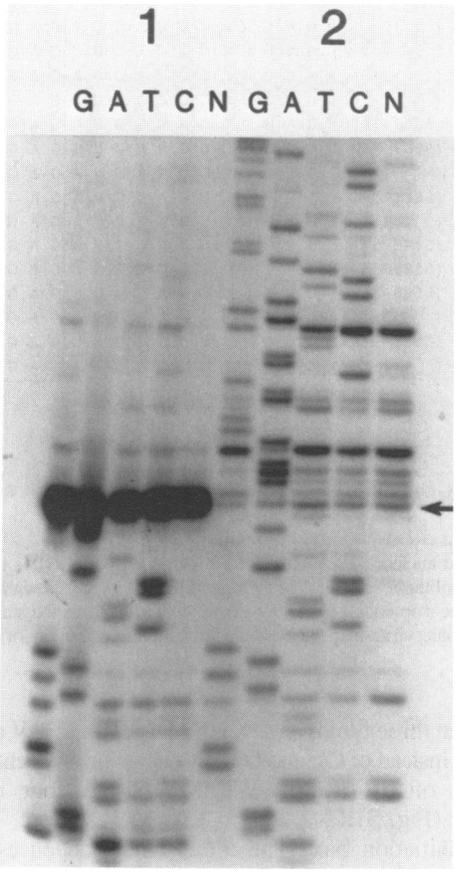

Fig. 2 Autoradiograph from the 5′ terminal sequence analysis of the subgenomic RNA of OYMV-Tin. The template RNA was extracted from either top component particles (1) or bottom component particles (2). G, A, T, C, N represent single sequencing reactions to which ddGTP, ddATP, ddTTP, ddCTP or no dideoxynucleotide was added respectively. The arrow indicates the first nucleotide of the subgenomic RNA.

not specifically terminated due to the presence of ddNTP but are the full-length transcripts of the subgenomic RNA which exists in a small amounts in the RNA preparation. Thus the 5′-terminal sequence of the OYMV-Tin subgenomic RNA was identified as 5′-AAU...-3′. The 5′-terminal sequence of KYMV-JB subgenomic RNA was deduced in the same way, and found to be the same. Thus the subgenomic RNAs of OYMV-Tin and KYMV-JB are similar to that of TYMV type strain (18) in that they all have the 5′ terminal sequence AAU... and are identical to the 3′ portions of the genomic RNAs from which they are transcribed.

Sequence similarities of the junction regions of tymoviruses

A comparison was made of the sequences of the region immediately to the 5′ side of the start codon of the virion protein gene of all tymoviral genomes, whose nucleotide sequences have been recently determined. As shown in Fig. 3, there is clear sequence homology in a region that is approximately 40 nucleotides long. In particular, two sequence blocks are closely similar in all tymoviruses sequenced so far:

(i) the tymobox. Seven or eight nucleotides to the 5′ side of the initiation site of the subgenomic RNA, there is a sequence block of sixteen nucleotides long that is shared by eleven tymoviruses. This tymobox sequence has only a single nucleotide

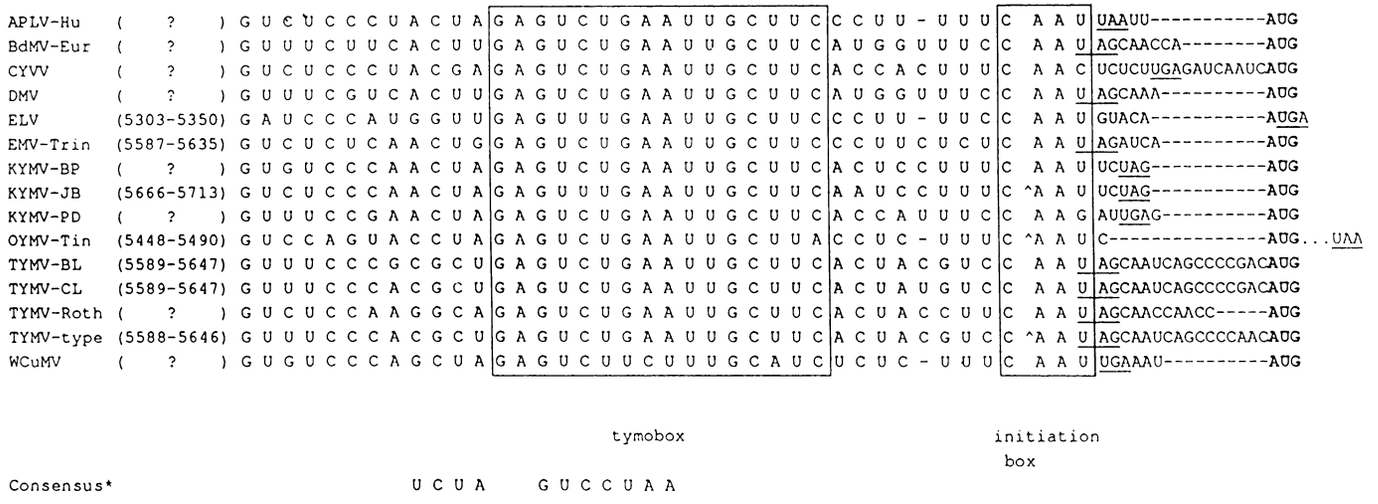

Fig. 3 Aligned nucleotide sequences of tymoviral genomic RNAs in a region surrounding the initiation site (indicated by arrows) of subgenomic RNA transcription. The positions of these sequences in their genomic RNAs are shown in brackets where complete genomic sequences have been determined. The two conserved sequences are boxed. The stop codon of the RP gene is underlined and the start codon of the VP gene is boldfaced. The consensus sequence is that of the subgenomic promoter of brome mosaic virus and the possible subgenomic promoter of alphaviruses (23).

difference in three tymoviruses, KYMV-JB and ELV at the fifth position (U instead of C), and OYMV-Tin at the sixteenth position (A instead of C), and in WCuMV it has four nucleotide differences (Fig. 3).

(ii) the initiation box. The 5′ terminal sequences of three tymoviral subgenomic RNAs (TYMV-type, OYMV-Tin and KYMV-JB) have been sequenced and found to start with the sequence AAU, and this triplet is preceded in the genome sequences of all three with a C. Indeed this sequence, CAAU, is found in all tymoviruses listed except in CYVV and KYMV-PD, where it has a C or G in the fourth position.

Tymovirus-specific probe

In dot-blot hybridization experiments, RNAs of APLV, ELV, BdMV-Eur, EMV- Trin, TYMV-CL, DMV, OYMV-Tin, and three isolates of KYMV (JB, W, PD) gave clearly positive signals, whereas RNAs of TMV and WCuMV gave little or no signal (Fig. 4), when probed with ^{32}P 5′ end-labelled oligonucleotide P3, which is complementary to the conserved tymobox sequence; it has been shown earlier that the tymobox sequence of WCuMV differs by four nucleotides from that of the consensus sequence. In another test the P3 probe hybridized strongly with genomic RNA of CoYMV, but not with RNAs of tobacco mosaic tobamovirus (type strain), tobacco mild green mosaic tobamovirus (U2 and U5 strains), sunnhemp mosaic tobamovirus, cucumber green mottle mosaic tobamovirus, beet western yellows luteovirus, barley yellow dwarf luteovirus (strain PAV), potato leafroll luteovirus and alfalfa mosaic virus, nor with nucleic acids extracted from virus-free *Brassica campestris* ssp. *pekinensis, Pisum sativum, Nicotiana glutinosa* and *N. tabacum* cv SR1. Thus the P3 probe detects, seemingly specifically, all tymoviruses tested except WCuMV; this exception provides a further measure of the specificity of the probe in addition to that obtained using genomic RNAs of non-tymoviruses.

A computer search of the EMBL (1989, release 19) and GENBANK (1987, release 48) nucleotide databases found no sequence among other viral and plant nucleotide sequences that have the same or greater sequence similarity with the tymobox sequence than the WCuMV tymobox sequence (12/16).

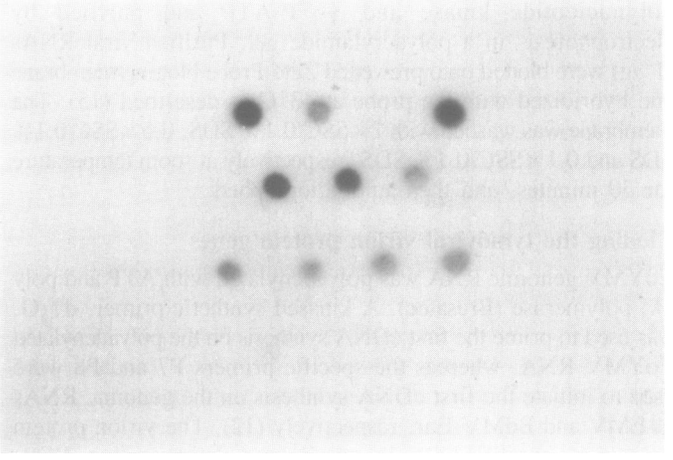

Fig. 4 Autoradiograph showing the DNA-RNA hybridization signal obtained using, as probe, the P3 oligonucleotide with sequence complementary to the tymobox. X-ray film was exposed to the membrane overnight at room temperature. Top row (from left to right): 1, APLV; 2, ELV; 3, WCuMV; 4, BdMV-Eur; Second row: 1, KYMV-JB; 2, KYMV-BP; 3, KYMV-PD; Third row: 1, EMV-Trin; 2, TYMV-CL; 3, DMV; 4, OYMV-Tin; Bottom row (between EMV and TYMV): TMV.

Cleavage of the RNA transcripts of three tymoviruses by the tymoribozyme

Fig. 3 shows that most of the known tymobox sequences include the sequence GUC, though in two it is GUU. The trinucleotide sequence GUX, where X is C or U or A, together with flanking sequences, that provide specificity, is the target of ribozymes derived from satellite RNAs of plant viruses (14). Thus a single ribozyme designed for the tymobox GUX should cleave all known tymoviral genomes except, probably, WCuMV which has four nucleotide substitutions in the tymobox sequence. Based on the

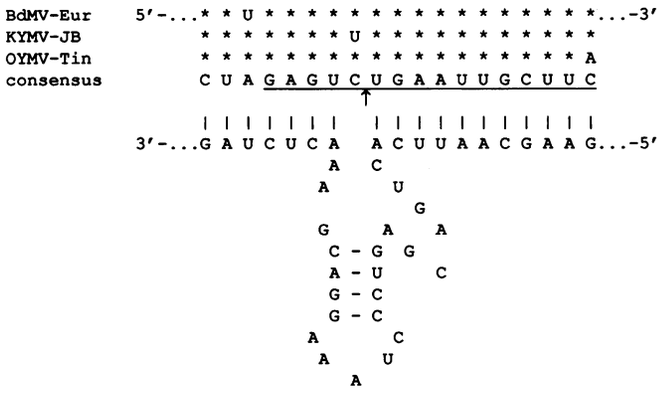

Fig. 5 The target sequences for the tymoribozyme in three substrate RNA molecules, also the sequence and folding of the tymoribozyme and its designed interaction with the targeted sequences.

* represents each nucleotide that is identical to that in the consensus sequence. The arrow shows the ribozyme cleavage site in these substrates. The tymobox sequence is underlined. The ribozyme catalytic domain is that of the subterranean clover mottle virusoid (J. Haseloff, personal communication).

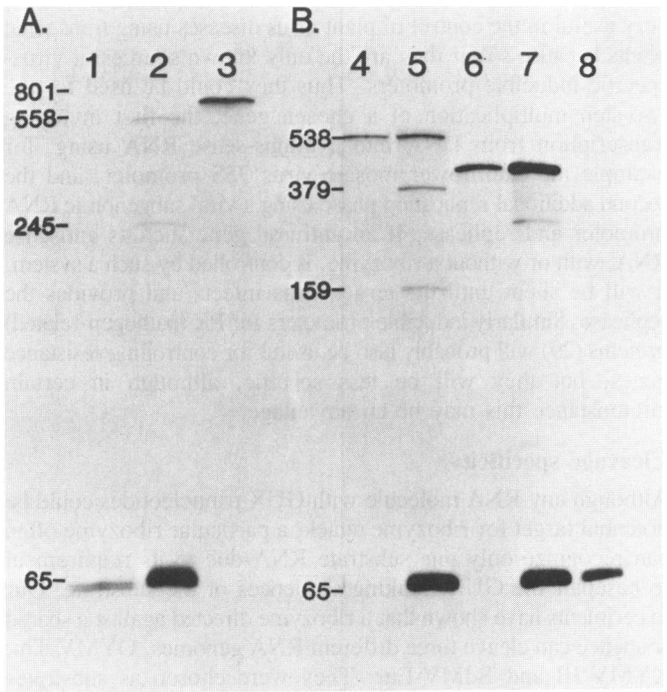

Fig. 6(A, B) Autoradiographs showing the results of in vitro treatment of OYMV-Tin (lanes 2–3), KYMV-JB (lanes 4–5) and BdMV-Eur (lanes 6–7) RNA transcripts by the tymoribozyme. Lanes 1 and 8, lane 3, lane 4 and lane 6 are the transcripts of plasmids pTzRz pOY41, pJB249 and pBM162 respectively. After incubation with the tymoribozyme, partial transcripts of OYMV-Tin (lane 2), KYMV-JB (lane 5) and BdMV-Eur (lane 7) were cleaved while the ribozyme remained unchanged. There was no cleavage when the tymoribozyme (lane 1 and 8), or transcripts of OYMV-Tin (lane 3), KYMV-JB (lane 4) and BdMV-Eur (lane 6) were incubated alone under the same conditions.

work of Haseloff and Gerlach (14), a 'tymoribozyme', was designed as shown in Fig. 5. The catalytic domain of the tymoribozyme was that of the subterranean clover mottle virusoid, which was shown to cleave RNA more efficiently than that of tobacco ringspot satellite RNA (J. Haseloff, personal

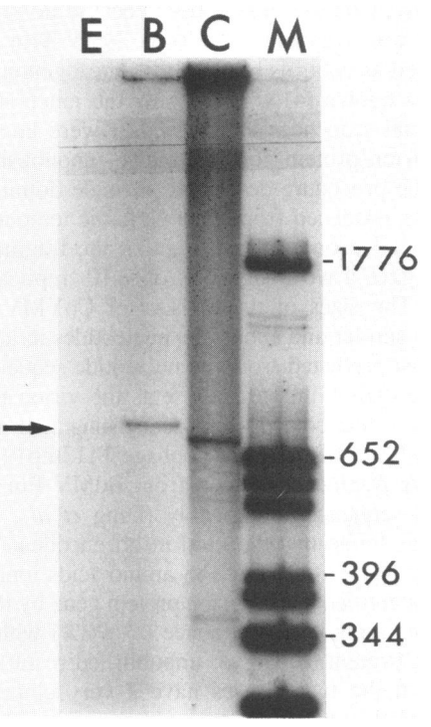

Fig. 7 PCR amplification of three virion protein genes: CoYMV (lane C), BdMV-Eur (lane B) and EMV-Trin (lane E). Lane M contains DNA markers (phage SPPI digested by EcoRI). The arrow points positions of dsDNA fragments representing the virion protein genes. PCR procedure (10 cycles): 94°C for 30 seconds; 46°C for 30 seconds and 70°C for 2 minutes.

communication). The number of nucleotides of the tymoribozyme that could basepair specifically with the flanking regions of the substrates were maximised, and included eleven-nucleotides (10 for OYMV) to the 3′ side and seven (4–6 in some) to the 5′ side of the GUX sequence (Fig. 5).

To test the activity of the tymoribozyme, the RNA molecules of the ribozyme and the three substrates, OYMV-Tin, KYMV-JB and BdMV-Eur, were transcribed from DNA clones pTyRz, pOY41, pKYJB249 and pBM162, respectively, and then purified by polyacrylamide gel electrophoresis. The length of the ribozyme, and of the fragments of OYMV, KYMV-JB and BdMV-Eur was 65, 838, 518 and 330 nucleotides approximately, including those nucleotides originating from the vector. The OYMV and KYMV transcripts were each mixed with the ribozyme in a 1:2 molar ratio at 50°C for one hour and this yielded fragments of 551 and 287 (Fig. 6A, lane 2) and of 360 and 158 nucleotides (Fig. 6B, lane 5) respectively. Cleavage also occurred, as predicted, when the ribozyme was incubated with BdMV-Eur RNA transcripts (Fig. 6B, lane 7). The lengths of the fragments produced by incubating with the ribozyme indicate that the cleavage occurred at the expected position, namely at the GUX tymobox site of those substrates. The ribozyme or viral transcripts when incubated alone, in the same conditions as the mixtures, remained unchanged in length (Fig. 6A and B).

Cloning of tymovirion protein genes

Because the tymobox is located to the 5′ side of the virion protein gene, it should be possible to use it in the genomic sense as a primer for initiating the synthesis of second-strand DNA complementary to the first-strand cDNA that had been synthesized using primer dT_8dG and polyadenylated tymoviral RNA. To test

this possibility, CoYMV was chosen because no sequence data on CoYMV was available at the time. EMV-Trin and BdMV-Eur were used as controls because the entire genomic sequence of EMV was known (4) whereas only the junction region and the 3′ terminal sequences of BdMV-Eur were known, but not the entire virion protein gene (J. Howe, unpublished results). Following the procedure described, a single dominant dsDNA fragment was produced from each genomic template using the same P4 as the second primer (Fig. 7); the fragment obtained from EMV gave a weak signal as only 10 'replication' cycles were used. The sizes of the cDNAs of CoYMV, EMV and BdMV were similar and about 700 nucleotides in length, which is close to that predicted from the nucleotide sequence of EMV (4). That the cDNA indeed represents the virion protein gene of EMV was further confirmed by sequencing after it was cloned into the SmaI site of the bacteriophage M13mp19.

The cDNA fragment obtained from BdMV-Eur was cloned in M13 and sequenced completely (Ding *et al.*, unpublished result). It was found that the viral insert encodes a single long open-reading frame (ORF) of 188 amino acids long. The ORF was further identified as the virion protein gene by the similarity of its encoded amino acid sequence (> 90%) with that of the DMV virion protein (J. Howe, unpublished result), because it is known that the two viruses have a very close serological relationship (19) and furthermore, with tymoviruses, serological relationship and virion protein sequence identity are significant correlated (20).

Using a similar approach, the cDNA of CoYMV was cloned and sequenced (Ding *et al.*, unpublished result) in M13 clone which was also shown to represent the virion protein gene by its close sequence similarity to that of the virion protein with CYVV (A. Mackenzie, unpublished result), with which it has a close serological relationship (19).

The technique described above is obviously useful, not only for providing useful taxonomic information, but could also be used to clone a tymovirion protein gene for transforming plants to attempt to render them resistant to infection (21).

DISCUSSION

Possible function of the junction region

The region of the genome, that includes the tymobox and initiation box, and which we call the 'junction region', probably functions as the promoter for subgenomic RNA transcription of tymoviruses. This seems likely because it occurs close to the start of the virion protein gene, and because it is the most conserved part of the tymoviral genomes sequenced so far. This sequence conservation seems not to reflect a need to conserve the amino acid sequence it encodes, because, firstly, the conserved domains of tymoviral replicase proteins have more nucleotide differences in third positions of codons, and secondly the nucleotide change in the fourth position of the tymoboxes of KYMV-JB and ELV leads to a change of the encoded amino acids. Thus it is the nucleotide sequence of the tymobox that is conserved rather than the amino acids it encodes.

Another reason for concluding that the tymobox is concerned with the virion protein mRNA production is that all other members of the supergroup of Sindbis-like viruses produce subgenomic RNA, a strategy for differentially expressing internal genes (22). A recent report (23) claimed that tymoviruses utilize the same mechanism as brome mosaic virus (BMV, 24) and alfalfa mosaic virus (AlMV, 25) to produce its subgenomic RNA, namely, by internal initiation of transcription on its minus-strand template. It has been clearly shown that, in both BMV (26, 27) and AlMV (25), there is a promoter element with a length less than 100 nucleotides that is responsible for initiating the subgenomic RNA transcription. The tymobox is located to the 5′ side of the start of the virion protein gene in the same position as the subgenomic RNA promoters of BMV and AlMV. Moreover, the tymobox and its surrounding sequences have some sequence similarity with the promoter sequences of BMV as well as alphaviruses (Fig. 3, and ref. 26). The genomes of several alphaviruses have a block of 21 nucleotides (including the first two nucleotides of the subgenomic RNA and 19 nucleotides preceding it) that is conserved. One difference is that in tymoviruses, the most conserved sequence is separated into two regions by a more variable region of 7 or 8 nucleotides, whereas in alphaviruses there is only one region that probably combines the functions of the tymobox and initiation box (28).

The separated conserved regions of the tymoviruses may have different functions. The tymobox probably acts as a signal sequence to be recognized and/or bound by viral RNA replicase whereas the initiation box may help to determine a precise start of the subgenomic RNA transcription.

Subgenomic promoters of RNA viruses will probably become very useful in the control of plant virus diseases using transgenic plants because so far they are the only known sources of virus-specific inducible promoters. Thus they could be used for the two-step multiplication of a chosen gene; the first involving transcription from DNA into a minus-sense RNA using, for example the cauliflower mosaic virus 35S promoter, and the second additional replication phase using a viral subgenomic RNA promoter and replicase. If an antiviral gene such as antisense RNA, with or without a ribozyme, is controlled by such a system, it will be silent until the target virus infects and provides the replicase. Similarly inducible promoters for PR (pathogen-related) proteins (29) will probably also be useful for controlling resistance genes, but they will be less specific, although in certain circumstance this may be an advantage.

Cleavage specificity

Although any RNA molecule with GUX trinucleotides could be potential target for ribozyme attack, a particular ribozyme often can recognize only one substrate RNA due to its requirement to basepair the GUX flanking sequences of the substrate. Our experiments have shown that a ribozyme directed against a shared sequence can cleave three different RNA genomes, OYMV-Tin, KYMV-JB and BdMV-Eur. They were chosen as substrates because the nucleotide sequences of the tymobox of these three viral genomes represent most of the known variation of this region among different tymoviruses (Fig. 3): a C ⇒ U transition in the fifth position of the KYMV-JB and ELV tymoboxes and a C ⇒ A transversion in the sixteenth position of the OYMV-Tin tymobox. Thus it is likely that this one ribozyme would cleave, equally well, most other tymoviral genomes except that of WCuMV.

The sequence of nucleotides at the cleavage site of naturally self-cleaving RNAs is highly conserved and usually GUC (30). In transcripts of newt satellite DNA, it was shown that cleavage activity remained essentially unaltered when the normal sequence was replaced by GUA or GUU, but was greatly diminished by GUG (31). Our experiments support these conclusions as the transcripts of KYMV-JB had the cleavage sequence GUU and were cleaved as well as those of the OYMV-Tin and BdMV-Eur transcripts which had the cleavage site GUC. Moreover the

ribozyme 'recognition' sequences were not completely homologous to two of the chosen transcript RNA sequences; the third nucleotide from the 5′ end of the target sequence of the BdMV-Eur transcript was not complementary to that in the 3′ end of the ribozyme, and there was a similar single mis-match at the 3′ end of the chosen OYMV-Tin target (Fig. 5). The ribozyme used in these experiments has one or fewer mis-matches with nine of the other known tymobox sequences (Fig. 5) and thus should cleave all of these.

A possible artificial virus-resistance gene

A number of artificial viral resistance genes have been used to transform plants and protect them against viruses. These genes, which differ in efficiency, include satellite RNAs of plant viruses (32, 33), virion protein gene (22, 34) and antisense RNA of viral genes (35). One disadvantage of the use of these genes is that they are very specific. However, if the ribozyme tested here, when transformed into plants, cleaved tymoviral genomes *in vivo* as efficiently as it does *in vitro*, it would protect these plants against most members of a single virus group, the tymoviruses.

One of the characteristics of viruses with RNA genomes is their high mutation rate and their ability to maintain relatively uniform populations by strong selection (36). Nonetheless this mutational pool allows them to respond rapidly to change, even though in a limited way compared with diploid and eukaryotic organisms. Therefore, if ribozymes were used as artificial plant resistance genes against RNA viruses, it is important to use as the target, a sequence, such as the tymobox, which is evolutionarily stable. Progeny genomic RNAs with mutations in their tymoboxes might escape from the attack of the ribozyme, but they would probably not be viable as they would no longer be recognized by their own replicase and hence not be able to produce virion messenger. We are examining these possibilities.

Routine detection of tymoviruses

The oligonucleotide, P3, complementary to the conserved tymobox sequence was found to be a tymovirus-specific probe. Furthermore, a taxonomy of nine tymoviruses (Ding *et al.*, unpublished result) based on the sequence similarities of a region with 100 nucleotides in length that is immediately 5′ to the tymobox correlated closely with the taxonomy of tymoviruses derived from larger parts of the genomic sequence or other criteria. Thus, this probe can be used as a primer to distinguish between individual tymoviruses and their isolates by dideoxynucleotide sequencing of viral genomic RNA, and comparing the sequence with the same region of the other tymoviral genomes. Virus group-specific probes, either nucleic acid or antibody, may offer a great potential in agricultural practice, for example, for quarantine purpose or ecological studies of certain virus groups, and it was recently reported that the antibody raised against the conserved core sequence of virion protein subunits can recognise specifically most potyviruses, the first virus group-specific antibody ever described (37).

ACKNOWLEDGEMENTS

We thank Dr Jim Haseloff for his advice. We also thank Mr. Mark Gibbs and Gary Mayo for synthesizing the oligonucleotide primers.

Abbreviations: APLV, Andean potato latent; BdMV-Eur, belladonna mottle, European isolate; CYVV, clitoria yellow vein; CoYMV, cacao yellow mosaic; DMV, dulcamara mottle; ELV, erysimum latent; EMV-Trin, eggplant mosaic, Trinidad isolate; KYMV, kennedya yellow mosaic (isolates: Bawley Point (BP), Jervis Bay (JB) and Port Douglas (PD)); OYMV-Tin, ononis yellow mosaic, Tintagel isolate; TMV, tobacco mosaic; TYMV, turnip yellow mosaic (isolates: Blue Lake (BL), Club Lake (CL), Rothamsted (Roth)); WCuMV, wild cucumber mosaic.

REFERENCES

1. Matthews, R.E.F. (1982). Intervirology 17, 138.
2. Morch, M.D., Boyer, J.C. and Haenni, A.L. (1988). Nucl. Acid Res. 16, 6157–6173.
3. Keese, P., Mackenzie, A., and Gibbs, A. (1989). Virol. 172, 536–546.
4. Osorio-Keese, M.E., Keese, P., and Gibbs, A. (1989). Virol. 172, 547- 554.
5. Ding, S., Keese, P. and Gibbs, A. (1989). Virol. 172, 555–563.
6. Ding, S., Gibbs, A. and Keese, P. (1990). J. gen. Virol. In press.
7. Pleij, C.W.A., Neeleman, A., van Vloten-Doting, L., and Bosch, L. (1976). Proc. Natl. Acad. Sci. USA 73, 4437–4441.
8. Guy, P.L., Dale, L.J., Adena, M.A., and Gibbs, A. (1984). Plant Pathol. 33, 337–346.
9. Blok, J., Gibbs, A., and Mackenzie, A. (1987). Arch. Virol. 96, 225–240.
10. Ou, J.-H., Strauss, E.G., and Strauss, J.H. (1981). Virol. 109, 281–289.
11. Maniatis, T., Frotsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
12. Gubler, U. and Hoffman, B.J. (1983). Gene 25, 263–269.
13. Melton, D.A., Krieg, P.A. Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. (1984). Nucleic Acid Res. 12, 7035–7056.
14. Haseloff, J. and Gerlach, W.L. (1988). Nature 334, 585–591.
15. Thomas, P.S. (1980). Proc. Natl. Acad. Sci. USA 77, 5201–5205.
16. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., and Erlich, H.A. (1988). Science 230, 1350–1353.
17. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., and Roe, B.A. (1980). J. Mol. Biol. 143, 161–178.
18. Guilley, H. and Briand, J.P. (1978). Cell 15, 113–122.
19. Koenig, R. (1976). Virol. 72, 1–5.
20. Ding, S. (1989). PhD Thesis. Australian National University. Australia.
21. Powell Abel, P., Nelson, R. S., De, B., Hoffmann, N., Rogers, S.G., Fraley, R.T., and Beachy, R.N. (1986). Science 232, 738–743.
22. Goldbach, R. (1987). Microbiol. Sci. 4, 197–202.
23. Gargouri, R., Joshi, R.L., Bol, J.F., Astier-Manifacier, S., and Haenni, A.L. (1989). Virol. 171, 386–393.
24. Miller., W.A., Dreher, T.W., and Hall, T.C. (1985). Nature 313, 68–70.
25. Langereis, K. (1987). PhD Thesis. State University of Leiden, The Netherlands.
26. Marsh, L.E., Dreher, T.W., and Hall, T.C. (1988). Nucleic Acid Res. 16, 981–995.
27. French, R. and Ahlquist, P. (1988). J. Virol. 62, 2411–2420.
28. Ou, J.-H., Rice, C.M., Dalgarno, L., Strauss, E.G., and Strauss, J.H. (1982). Proc. Natl. Acad. Sci. USA 79, 5235–5239.
29. Bol, J.F., Hooft van Huijsduijnen, R.A.M., Cornelissen, B.J.C., and van Kan, J.A.L. (1987). In: Plant Resistance to Viruses (Evered, D. and Harnett, S. eds). pp72–91. Ciba Found Symp. 133. John Wiley & Sons, Chichester.
30. Forster, A.C. and Symons, R.H. (1987). Cell 49, 211–220.
31. Koizumi, M., Iwai, S., and Ohtsuka, E. (1988). FEBS Lett. 228, 228–230.
32. Baulcombe, D.C., Saunders, G.R., Bevan, M.W., Mayo, M.A., and Harrison, B.D. (1986). Nature 321, 446–449.
33. Gerlach, W.L., Llewellyn, D., and Haseloff, J. (1987). Nature 328, 802–805.
34. Turner, N.E., O'Connell, K.M., Nelson, R.S., Sanders, P.R., Beachy, R.N., Fraley, R.T., and Shah, D.M. (1987). EMBO J. 6, 1181–1188.
35. Baulcombe, D.C., Hamilton, W.D.O., Mayo, M.A., and Harrison, B.D. (1987). In: Plant Resistance to Viruses (Evered, D. and Harnett, S. eds). pp170–178. Ciba Found. Symp. John Wiley & Sons, Chichester.
36. Steinhauer, D.A. and Holland, J.J. (1987). Ann. Rev. Microbiol. 41, 409–433.
37. Shukla, D.D. and Ward, C.W. (1989). Adv. Virus Res., 36, 273–314.