Anti-sense regions in satellite RNA of cucumber mosaic virus form stable complexes with the viral coat protein gene

M.Ali Rezaian and Robert H.Symons

CSIRO Division of Horticultural Research, GPO Box 350, Adelaide, SA 5001 and Adelaide University Centre for Gene Technology, Department of Biochemistry, University of Adelaide, SA 5000, Australia

Received 31 January 1986; Accepted 19 March 1986

ABSTRACT

The interaction in vitro of the RNA of the Q-strain of cucumber mosaic virus (CMV) with its satellite RNA (sat-RNA) has been studied. In hybridisation reactions containing 30% formamide at 45°C, sat-RNA binds to CMV RNA 3 and 4 but not to CMV RNA 1 and 2 or RNA from tobacco mosaic virus and alfalfa mosaic virus. The viral coat protein gene present in RNA 3 and 4 contains the site of binding but this region does not contain complementary sequences of any significant length to the sat-RNA sequence. However, the optimum alignment of short complementary sequences present in these regions revealed a stable structure in which it is proposed that sat-RNA twists around the coat protein gene so that two separate blocks of nucleotides in sat-RNA base pair in opposite directions with two adjacent blocks in the coat protein gene to form a knot-like structure. The binding site is a region of 33 nucleotides within the coding region of the coat protein gene which base pairs with residues 98-113 and 134-152 of sat-RNA. The possibility of the binding region of sat-RNA functioning as an "anti-sense" sequence in regulation of the viral coat protein synthesis is discussed.

INTRODUCTION

The genome of cucumber mosaic virus (CMV) consists of three single-stranded RNAs designated RNA 1 (3389 nucleotides) (1), RNA 2 (3035 nucleotides) (2), and RNA 3 (2193 nucleotides) (3). A fourth RNA is also encapsidated which is subgenomic and its entire sequence of 1027 nucleotides containing the viral coat protein gene is present in the 3'-region of RNA 3 (3).

A small single-stranded RNA of 334 to 339 nucleotides is also found in some isolates of CMV (4,5). This RNA cannot replicate in the absence of CMV and requires CMV functions for its replication. It is therefore known as CMV-satellite RNA (sat-RNA). sat-RNA has also been termed CMV associated RNA 5 or CARNA 5 (6). The nucleotide sequence of sat-RNA (7) was compared with the sequences of the viral genome both by cDNA hybridisation in liquid (8) and by computer analysis of their entire nucleotide sequences (1) and no extended regions of homology was detected. However, a region of nucleotides at the 5'-end of sat-RNA was found to be complementary to a part of the non-coding sequences at the 5'-end of CMV RNA 1 and 2. These complementary regions in both sat-RNA and the viral RNAs are parts of similar hair-pin loop structures (1).

The presence of complementary sequences and similar secondary structures in sat-CMV and CMV RNA 1 and 2 indicate the possibility of specific interaction between CMV and its satellite. It was therefore desirable to test whether sat-RNA would bind to CMV RNA in vitro. Results reported here demonstrate that, in vitro, sat-RNA does not bind strongly to CMV RNA 1 and 2 as may be predicted but specifically interacts with the CMV coat protein gene by forming an unconventional knot-like structure.

**MATERIALS AND METHODS**

**Preparation of Viral RNA and sat-RNA**

The Q-strain of CMV (9) was grown (10) and the viral RNA was isolated from purified virus (11). Individual viral RNAs and sat-RNA were purified by polyacrylamide gel electrophoresis (12). The genomic CMV RNAs used in this study were from virus preparations which did not contain the satellite.

**5'-32P Labelling of sat-RNA**

The m⁷G cap structure of sat-RNA was removed chemically (13). Ten ug of decapped RNA was incubated with 0.05 unit of calf intestinal phosphatase in a volume of 10 ul containing 0.1M Tris, pH 8, at room temperature for 30 min. The phosphatase, prior to use, was treated with diethyl pyrocarbonate to eliminate contaminating nucleases (14). The phosphatase was inactivated by adding 2 ul of 30mM nitrioloacetic acid, pH 7.6, and incubating at room temperature for 20 min followed by heating at 100° for 1 min (15). One ug of decapped sat-RNA was labelled at the 5'-terminus (3) using γ-³²P-ATP (BRESA, Adelaide) and polynucleotide kinase (Boehringer). The reaction mixture was applied to a 6% polyacrylamide gel containing Tris-borate-EDTA and 7M urea (16) and the sat-RNA was eluted from the single radioactive band in the gel (17) detected by autoradiography.

**Nucleic Acid Hybridisation**

Hybridisations were carried out as previously described in (18). Reactions in a volume of 20 ul contained 2 x SSC (0.3M NaCl, 0.03M tri-sodium citrate, pH 7.5), 0.5% sodium dodecyl sulphate, 30% formamide (unless otherwise stated) and the indicated amounts of viral and sat-RNAs. The mixture was heated at 85° for 3 min followed by incubation at 45° for 1 hr. Three ul of 80% (v/v) glycerol containing 0.02% bromophenol blue was added before electrophoresis in 1.2% agarose gels in Tris-borate-EDTA (16) at 100V at room temperature using a Pharmacia GNA100 system. Gels were stained with ethidium bromide and photographed. Gels containing ³²P-sat-RNA were dried with a hair dryer for a period of 45 min and autoradiographed.

**RESULTS**

**sat-RNA Binds to the Helper Viral RNA**

From the comparisons of sequence data, it was known that there were no extended
regions of complementary sequences between sat-RNA and the Q-CMV RNAs (1). To examine the base pairing of short complementary regions such as those found at the 5'-end of sat-RNA and CMV RNA 1 and 2 it was necessary to carry out hybridisation at an optimum stringency. In an initial experiment, 5'-32P-labelled sat-RNA was incubated with total CMV RNA under hybridisation conditions in the presence of an increasing concentration of formamide. The RNA was then analysed by electrophoresis in agarose. Results shown in Fig. 1A demonstrated that, following hybridisation, CMV RNAs remained intact as detected by staining with ethidium bromide, although a small proportion migrated more slowly than RNA 1 when incubated in 20% or 30% formamide (Fig. 1A). A minor band migrating between RNA 3 and 4 was detected when incubations were in 20% formamide or a higher concentration (Fig. 1A); this species was not further characterised. An autoradiograph of the same gel is shown in Fig. 1B which demonstrated that, following hybridisation, a proportion of the 32P-labelled sat-RNA migrated as multiple bands with slower mobilities. There were no radioactive bands corresponding to RNA 1 and 2 but, at 10 and 30% formamide and above, two of the radioactive bands had mobilities just slower than RNAs 4 and 3, respectively. These two
bands were not present in hybridisation products in the absence of formamide (Fig. 1B). A major radioactive band appeared in gels between RNA 4 and sat-RNA (Fig. 1B) especially when hybridisation was done in 20% or lower concentrations of formamide. There was no visible RNA species corresponding to this band in the stained gel (cf. Fig. 1A and B). The origin of nucleic acid generating this radioactive band was not further investigated. The electrophoretic mobility of sat-RNA incubated under the hybridisation conditions in the presence or absence of various concentrations of formamide remained unchanged (results not shown, see next section).

sat-RNA Binds to the Viral Coat Protein Gene

Using 30% formamide as an optimum concentration the specificity of binding of sat-RNA to CMV RNAs was examined. RNAs from alfalfa mosaic virus (AMV) and tobacco mosaic virus (TMV) were used as controls and were found not to hybridise with sat-RNA (Fig. 2A). In the absence of any viral RNA the electrophoretic mobility of sat-RNA did not change upon incubation under hybridisation conditions. The multiple bands observed after hybridisation of sat-RNA to CMV RNA (Fig. 2A) are not, therefore, due to self-annealing of sat-RNA (Fig. 2A).

In order to determine which of the CMV RNA species were involved in binding with the satellite, the four CMV RNAs were used separately in hybridisation reactions with $^{32}$P-labelled sat-RNA. Figure 2B shows that only RNA 3 and 4 significantly bound to sat-RNA. The presence of unlabelled sat-RNA in the same experiment, out-competed $^{32}$P-sat-RNA binding to CMV RNA (Fig. 2B). The radioactive hybrid products had slightly slower mobilities than the respective RNA 3 and 4 bands detected by
FIGURE 3. Gel electrophoresis of CMV RNAs hybridised with unlabelled sat-RNA as detected by staining with ethidium bromide. Three ug of total CMV RNA was hybridised with 1 ug unlabelled sat-RNA and analysed in 1.2% agarose gel (A). In (B) 1 ug each sat-RNA and RNA 4 were used and electrophoresis was done in a 2% agarose gel. Additional RNA bands detected following hybridisation of CMV RNA and sat-RNA are identified by arrows.

staining (Fig. 2B). In fact, when stainable quantities of unlabelled sat-RNA were hybridised with total CMV RNA, followed by electrophoresis, RNA 3 and 4 each migrated as double bands, one with the mobility of RNA 3 or 4 and the other with a slower mobility (Fig. 3A). When the hybrid of RNA 4 and sat-RNA was electrophoresed in 2% agarose the double band was clearly resolved (Fig. 3B). Slower electrophoretic mobilities are expected if sat-RNA binds to RNA 3 and 4. Although the entire sequence of RNA 4 is present in RNA 3, the possibility that sat-RNA may also bind to the non-RNA 4 part of RNA 3 cannot be eliminated.

These results were unexpected because the nucleotide sequence of sat-RNA does not contain any significant blocks of nucleotides which are complementary to Q-CMV RNA sequences and yet sat-RNA binds to CMV RNA 3 and 4 under stringent conditions. Moreover, sat-RNA did not bind to RNA 1 and 2 which contain the 5'-complementary sequence of 18 nucleotides with a single residue mismatch and a bulge of one nucleotide (1).

Site of Binding of sat-RNA to the Viral Coat Protein Gene

In the absence of extended regions of complementary sequences in sat-RNA and in RNA 4, the role of short complementary sequences in the observed binding was
TABLE 1: A list of short complementary regions found between Q-CMV-RNA 4 and sat-RNA.

<table>
<thead>
<tr>
<th>Number of base pairs in the region</th>
<th>Number of mismatched residues</th>
<th>Number of G:U base pairs</th>
<th>Number of single residue bulges</th>
<th>Number of occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15*</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14*</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Complementary sequences were searched using a computer program based on an algorithm of Goad and Kanehisa (19). Of the two neighboring sequences used in Fig. 4, one contains 15 and the other 14 complementary base pairs and these are shown by asterisks.

Examined. Table 1 lists all regions in the nucleotide sequences of sat-RNA and Q-CMV RNA 4 which are complementary as detected by computer analysis. None of the complementary sequences appear singularly capable of stable base pairing in the presence of up to 50% formamide under the hybridisation conditions used. For example, complementary blocks of 18 oligodeoxynucleotides with 11 G:C residues are not able to form stable hybrids in 30% formamide in otherwise similar conditions (18) to those used in the present study. The locations of complementary blocks in sat-RNA and in RNA 4 are dispersed, except for one site in RNA 4 where two complementary blocks to sat-RNA are adjacent. The corresponding blocks in sat-RNA are separated by 20 nucleotide residues and are not co-linear with those in RNA 4 as illustrated in Fig. 4A. The two blocks of sequences in RNA 4 are from residues 557 to 585 within the coat protein gene.

Base pairing of RNA 4 and sat-RNA involving these two blocks is possible if sat-RNA twists around RNA 4 to form a structure shown in Fig. 4B. This proposed knot-like structure appears quite stable because it involves a continuous stretch of 33 residues on the RNA 4 including 17 G-C base pairs. There are 3 single residue mismatches, 1 two residue mismatch, 1 G-U pairing and no bulges in the structure. Residues 571 and 572 of RNA 4 shown in Fig. 4B can pair with either of the two complementary segments in sat-RNA, presumably making the structure more stable at the ends. Such stability is compatible with the binding of sat-RNA to RNA 4 in high stringency conditions. Alignment of all other complementary sequences shown in Table 1 did not produce any other structure in which complementary sequences would overlap.
FIGURE 4. Proposed binding regions of Q-CMV-RNA 4 and sat-RNA. A schematic representation of two complementary regions between these RNAs is shown in (A). Broken lines in (A) demonstrate the antiparallel location of region 1 and 2 relative to each other. RNA 4 (1027 residues) and sat-RNA (336 residues) are not drawn to scale. The proposed base pairing of residues in these two regions is shown in (B). The residue numbers of RNA 4 are from the 5'-end of this RNA as obtained from the RNA 3 sequence (3) and those of the sat-RNA sequence are from (7). The A:U base pair, bordered by mismatches, between residue 579 of RNA 4 and residue 144 of sat-RNA may not form in solution.

Base pairing of distant sequences within an RNA molecule has been predicted (20) and detected in viral RNAs (21) but interaction of RNA molecules by the structure shown in Fig. 4B is unconventional. It was therefore desirable to examine experimentally whether the region of sat-RNA and coat protein gene shown in Fig. 4B was the site of binding. A synthetic 33-mer oligodeoxynucleotide complementary to the entire segment of RNA 4 which is involved in the knot-like structure shown in Fig. 4B was prepared and used in competition hybridisation experiments. When 32P-labelled sat-RNA was used to hybridise to RNA 4 in the presence of the oligonucleotide, its hybridisation was almost completely inhibited (Fig. 5A). This experiment was also done with unlabelled sat-RNA (Fig. 5B). RNA 4 incubated with the complementary oligonucleotide had a slightly slower electrophoretic mobility as determined by ethidium bromide staining, confirming that the
FIGURE 5. Inhibition of sat-RNA binding to RNA 4 by a synthetic 33-mer oligodeoxynucleotide complementary to residues 557-589 of RNA 4 shown in Fig. 4B. 32P-labelled sat-RNA was hybridised with RNA 4 in the presence and absence of the 33-mer and the product was analysed by agarose gel electrophoresis and autoradiographed (A). A sample of sat RNA without RNA 4 was also incubated as a control in (A). In (B) unlabelled sat RNA 4 was used to identify the hybrid band from RNA 4 by staining with ethidium bromide.

33-mer was binding to it (Fig. 5B). In the presence of the 33-mer, the binding of sat-RNA to RNA 4 was inhibited but in its absence the hybrid band moving slower than RNA 4 in the gel was detected (Fig. 5B). This inhibition by the oligonucleotide is specific to the site of interaction of RNA 4 and sat-RNA because the presence of other sequences, contained in CMV RNAs 1 and 2, in the hybridisation reaction did not interfere with the observed binding (Figs. 1B, 2A, 2B, 3A). These results confirm that the regions of the coat protein gene and sat-RNA shown in Fig. 4B are involved in binding. It is conceivable that, once a stable binding of sat-RNA to RNA 4 has occurred, other regions of the two molecules could interact to produce a specific 3-dimensional configuration of enhanced stability.

DISCUSSION

In this paper we demonstrated that CMV satellite RNA specifically binds to residues 557-589 of the coat protein gene of Q-CMV RNA 4. The binding occurred at 45° in formamide concentrations as high as 50% in 2 x SSC. This is a stringent hybridisation condition and would not allow base pairing of short complementary sequences (18). The largest region of homology between sat-RNA and RNA 4 contains 23 base pairs with one single residue bulge and 4 mismatches (Table 1). It is clear that the binding observed in Fig. 1 was not due to conventional base pairing of any single pair of complementary RNA
Binding of sat-RNA to RNA 3 and 4 did not occur in the absence of formamide (Fig. 1B) and was inefficient at formamide concentrations below 20%. This could probably be due to the presence of secondary structures in the regions of CMV RNA and/or sat-RNA involved in binding. Secondary structures in the coding regions of CMV RNA have not been determined but a part of the binding site of sat-RNA shown in Fig. 4 (residues 99 to 113) has been shown to be involved in secondary structure by a site-specific enzymic cleavage method (7). At formamide concentrations below 30%, $^{32}$P-labelled sat-RNA efficiently bound to a nucleic acid species in total CMV RNA which had an electrophoretic mobility faster than RNA 4 (Figs. 1B, 2A, 2B). This nucleic acid could be one of the subgenomic RNAs which have been detected in CMV RNA (22).

In CMV RNA a strong block of complementary sequence to sat-RNA exists in a conserved region of the 5'-noncoding region of RNA 1 and 2 (1). It was therefore interesting to note that, in vitro, sat-RNA did not bind to RNA 1 and 2 but rather bound to RNA 3 and 4. Both in sat-RNA and the 5'-conserved region of RNA 1 and 2 the complementary sequences are parts of similar hairpin structures (1). Lack of hybridisation between sat-RNA and RNA 1 and 2 could be explained by the presence of these hairpins, the stability of which prevent any hybrid formation. It is possible that in the presence of excessive concentrations of sat-RNA the binding equilibrium may favour the hybridisation of some sat-RNA to the 5'-conserved region. Possible interaction of the 5'-structures in sat-RNA and RNA 1 and 2 have been discussed (1).

It appears that in the CMV plus satellite RNA system specific mechanisms have evolved for the replication and regulation of RNA components. The satellite RNA cannot replicate in the absence of CMV genomic RNAs in whole plants or isolated protoplasts (23), it shows a varying degree of specificity to various strains (24), it specifically reduces the level of RNA 1 and 2 synthesis (24), it has a small "anti-sense" sequence to a conserved region of RNA 1 and 2 (1), it forms a similar hairpin structure at the 5'-end as the structure in RNA 1 and 2 (1), it has a partial 3'-terminal homology to CMV genomal RNAs (7) and, finally, it specifically binds to the coat protein gene (Fig. 4B). In RNA 4, the proposed site of hybridisation with sat-RNA is located within the coat protein reading frame. By analogy with the closely related brome mosaic virus, the initiation site of CMV RNA replication would be in the noncoding region of the RNAs (25-27). Hence, it is feasible that the binding of sat-RNA to the coat protein gene could regulate coat protein synthesis. Inhibition of protein synthesis by the interaction of "anti-sense" RNAs with cellular mRNAs is well established (28-31). A similar phenomenon has been reported in RNA phage-infected E. coli (32). It appears, therefore, that in the cellular environment complementary sequences hybridise efficiently. sat-RNA may regulate viral coat protein synthesis by a similar mechanism.
Intramolecular interactions involving viral RNAs and satellite RNAs have not been reported before. A similarity to the finding presented here exists in adenovirus associated small RNAs (VA RNAs) which bind to the late viral mRNA. The possibility has been raised that this binding may be involved in mRNA splicing (33).

The proposed structure shown in Fig. 4B is unconventional in involving two RNA molecules in a knot-like structure. Secondary structures in which distant complementary sequences of an RNA molecule could form knot structures have been predicted (20). In such a structure one end of an RNA molecule passes through a loop structure and forms additional base pairs. Another type of secondary structure involving distant residues are referred to as pseudoknots in which parts of loops in two hairpins base pair so that distant double helices align in a coaxial manner (20). While pseudoknots are common in the tertiary structure of the tRNA-like structure at the 3'-end of several plant viral RNAs, true knots have not been reported (21, 34). The structure shown in Fig. 4 is distinct from either type of structure mentioned above because it brings together two RNA molecules but it is similar to pseudoknots of plant viral RNAs in aligning two short helices in coaxial form.

Plant viral satellites have a versatile mode of interaction with their helper viruses (4, 5). It will be of interest to examine whether other viral satellite RNAs bind to their respective viral genome RNAs.

ACKNOWLEDGEMENTS

We thank Dr. Derek Skingle for the synthetic oligonucleotide, Rhys Williams and Chris Davies for assistance in computer analysis and useful discussion, Jennifer Cassady and Susan Harris for assistance. This work was supported by The Australian Research Grants Scheme and by a Commonwealth Government Grant to the Adelaide University Centre for Gene Technology in the Department of Biochemistry.

REFERENCES