A small DNA molecule containing covalently-linked ribonucleotides originates from the large intergenic region of the cauliflower mosaic virus genome

Simon N. Covey*, David Turner and Gert Mulder

Department of Virus Research, John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

Received 19 October 1982; Revised and Accepted 13 December 1982

ABSTRACT

We have detected a small DNA molecule (sa-DNA), 725 nucleotides long, in cauliflower mosaic virus (CaMV)-infected, but not non-infected, turnip leaves. Alkali and RNase A treatments shortened sa-DNA by 100 nucleotides and we conclude that it contains covalently-linked ribonucleotides. This DNA co-purified with cellular polyadenylated RNA. It is complementary to the β-strand of CaMV DNA and of opposite polarity to RNAs transcribed from the α-strand. Hybridisation studies suggest that sa-DNA originates from the large intergenic region (IR1) of the CaMV genome. A small double-stranded DNA with three single-stranded components, which co-purifies with cellular DNA, appears to be related to sa-DNA but lacks detectable ribonucleotides.

INTRODUCTION

Studies of the double-stranded DNA genome of cauliflower mosaic virus (CaMV), have been principally concerned with the development of a vector for gene transfer in plants together with the closely associated goal of understanding the mechanisms which control the expression and replication of plant genes.

An analysis of the distribution of termination codons in the nucleotide sequence (ca. 8000bp) of CaMV DNA (1,2) has revealed six (to eight) (3) adjacent or overlapping open regions which possibly encode proteins. The open regions are interrupted by the large (IR1) and small (IR2) intergenic regions (3,4). The termini of two abundant viral polyadenylated RNAs are transcribed from within the intergenic regions (5,6). The 5'-end of the 1.9 Kb mRNA which encodes a 62,000 MW virus inclusion body-associated protein (7) is transcribed from IR2 eleven nucleotides upstream of open region VI (5). The 3'-end of this RNA maps in IR1 and is co-terminal with the 3'-end of 35S RNA, a transcript of the complete CaMV DNA α-strand. The structure of 35S RNA is unusual in that its 5'-terminus is located about 200 nucleotides upstream of its 3'-terminus (5)(See Fig. 1).

CaMV virion DNA contains three discontinuities (8) and one of these, G1, is in the α-strand, which is of opposite polarity to CaMV RNA

0305-1048/83/1102-0251$2.00/0

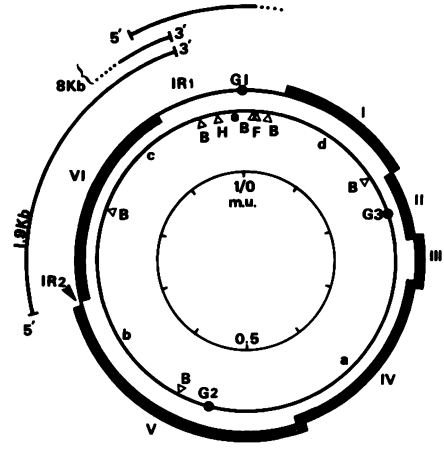

Figure 1. Map of the CaMV genome. The non-transcribed strand contains 2 discontinuities (G2, G3). The transcribed α-strand has a single discontinuity (G1) and shows the location of six long open regions (possible coding regions I-VI) which would appear in RNAs transcribed from the α-strand. The open regions are interrupted by the large (IR1) and the small (IR2) intergenic regions which contain the termini of the 1.9Kb and 8Kb (35S) RNA transcripts. Restriction enzyme sites relevant to this paper: H, Hinf 1; F, Bst EII; B, Bgl II and Bgl II fragments labelled a-e.

transcripts (9,10), in IR1. Since transcription of 35S RNA appears to traverse G1, it is likely that a covalently-closed circular form of CaMV DNA (11,12) and not virion DNA, is transcribed. It is not yet known how the discontinuities function although they may be related to DNA replication.

In this paper, we report the existence of small DNA molecules, isolated from virus-infected tissue, which originate from IR1 close to the α-strand gap position and the 35S RNA termini. One small DNA species contains covalently-linked ribonucleotides and appears to be associated with viral large RNA transcripts.

MATERIALS AND METHODS

Virus, DNA and recombinant DNA clones. Cauliflower mosaic virus isolate cabbage B-John Innes (cabb. B-JI) was propagated in turnip (Brassica rapa L. cv. Just Right). Virus was purified and virion DNA was extracted as described by Hull (13). Hybridisation probes were prepared from virion DNA and from the five major CaMV DNA Bgl II fragments cloned into pKC7, by R. Hull.

Isolation and fractionation of cellular nucleic acid. Cellular total

nucleic acid was extracted from CaMV-infected turnip leaves 18-22 days after inoculation, using the method described by Covey and Grierson (14). Total cellular nucleic acid was separated into a fraction which was soluble, and a fraction which was insoluble, in 3 M sodium acetate pH 6. Poly(A)+RNA was isolated from the salt-insoluble fraction using oligo (dT)-cellulose as previously described (5); a heat-denaturation step was omitted. Sucrose gradient ultracentrifugation was also as described previously (5).

Gel electrophoresis. Non-denaturing gels : nucleic acid samples were electrophoresed through 1.5% agarose slab gels in a buffer containing 25 mM Tris-acetate pH 7.9, 1 mM EDTA, 5 mM sodium acetate. Glyoxal gels : nucleic acid (approx 1-2μg) was glyoxalated by incubation in 1 M glyoxal (deionised), 70% formamide (deionised), 10 mM sodium phosphate buffer pH 7.0 at 55° for 15 min and then fractionated with 1.5% agarose gels in the same buffer used for non-denaturing gels. Two-dimensional gel electrophoresis was as described by Favaloro et al.(15).

Blotting and hybridisation. Gel-fractionated nucleic acid was blotted from neutral gels (16), glyoxal gels (17) and 2-D gels (15) onto nitrocellulose. Blots were probed by hybridisation with CaMV virion DNA or cloned CaMV Bgl II fragments labelled by nick-translation. Strand separation of nick-translated CaMV virion DNA was achieved by electrophoresis in alkaline agarose gels followed by elution of the appropriate bands.

Size markers for gel electrophoresis were cloned DNAs of the CaMV genome (8Kb); Bgl II fragments -a (3.3Kb), -b (1.9Kb), -c (1.1Kb), -d (1.0Kb); Eco Rl-d (0.45Kb) and Eco Rl-d digested with Hind III (0.25Kb and 0.2Kb).

Enzyme and alkali treatments. Samples were treated with DNase I in a mix containing 50 mM Tris-HCl pH 7.6, 5 mM magnesium chloride, 20 μg/ml DNase I incubated at 20° for 30 min. RNase treatment was in 0.3 M sodium chloride, 0.03 M trisodium citrate and nucleic acid samples were digested at 37° for 30 min with 30μg/ml of pancreatic RNase A. Following treatment with each enzyme, mixtures were extracted with phenol/chloroform (1:1) and undigested nucleic acid was precipitated with isopropanol, together with E. coli tRNA as carrier. Samples were treated with alkali in a mix containing 0.15 M sodium hydroxide, 10 mM EDTA incubated at 60° for 15 min. The mixture was neutralised and resistant material precipitated with isopropanol. Digestions with restriction enzymes Bgl II and Hinf 1 were under conditions recommended by the suppliers (BRL Inc.).

RESULTS

Distribution of CaMV sequences in cellular nucleic acid fractions.

Cellular nucleic acid, isolated from CaMV-infected turnip leaves, was fractionated by high-concentration salt to produce a soluble fraction, which contained RNA smaller than 5S and DNA, and an insoluble fraction, which contained mostly RNA larger than 5S. CaMV-specific sequences were detected in each fraction by agarose gel electrophoresis and blot-hybridisation of glyoxalated nucleic acid. Mixed populations of DNA and RNA were thus analysed under denaturing conditions and their sizes compared directly with CaMV DNA size markers.

CaMV-specific sequences present in the salt-soluble (DNA) fraction ranged in size from 0.2Kb to larger than 8Kb (Fig. 2a,b). Many of the bands were single-stranded components produced by denaturation of various cellular forms of CaMV DNA some of which we have characterised (Hull and Covey, in

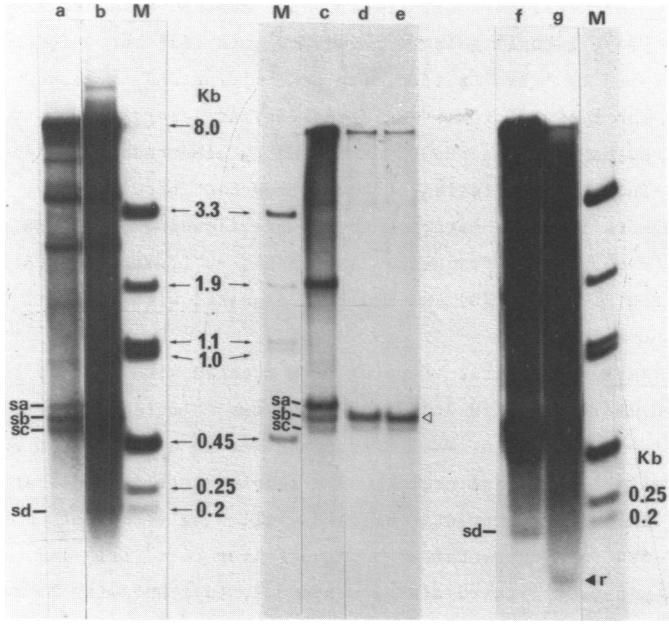

Figure 2. CaMV-specific sequences in turnip leaf nucleic acid fractions. Nucleic acid was glyoxalated and electrophoresed in agarose gels then blotted on to nitrocellulose and probed with nick-translated CaMV DNA. (a), the cellular DNA fraction with the film over-exposed in (b) to show sd-DNA. (c), the poly(A)+RNA fraction (over-exposed in [f] to show sd-DNA) treated with (d), alkali and (e), RNase A. Size markers (M) are CaMV DNA restriction enzyme fragments.

preparation). In addition, four small single-stranded nucleic acid species, with sizes of 725, 625, 575 (Fig. 2a) and 200 nucleotides (Fig. 2b) were detected. We have designated these sa-, sb-, sc- and sd- respectively.

CaMV virions contain very small quantities of sb-, sc- and sd-DNAs (but not sa-DNA). However, the amount of these small DNAs in purified virions is considerably less than that found in cellular DNA preparations from infected leaves. Co-extraction of nucleic acid from healthy leaves with purified CaMV releases very little virion DNA. Also, co-extraction of purified virion DNA with nucleic acid from healthly leaves does not increase the small amount of sb-, sc- and sd-DNAs normally observed in a virion DNA prepapration (data not shown). From this we conclude that sb-, sc- and sd-DNAs exist in infected cells largely as non-encapsidated molecules and that they do not arise as a result of breakage of virion DNA during extraction.

CaMV-specific sequences, including the four small nucleic acid species sa-, sb-, sc- (Fig. 2c) and sd-, (Fig. 2f), were also present in a polyadenylated RNA (poly(A)+RNA) fraction purified by oligo (dT)-cellulose separation of salt-insoluble RNA. However, the 725 nucleotide component (sa-) was significantly enriched in this fraction (Fig. 2c).

In our previous studies of CaMV polyadenylated transcripts (5,7) the standard preparative procedure included DNase treatment of salt-insoluble nucleic acid and heat-denaturation before oligo (dT)-cellulose fractionation. Under these conditions, we detect two major polyadenylated species : 35S RNA and the 1.9Kb mRNA encoding the virus inclusion body-associated protein and sometimes a minor species of 0.6Kb which are completely susceptible to hydrolysis by alkali (data not shown). Since the small nucleic acid species (sa- to sd-) found in the non-DNase-treated poly(A)+RNA preparation are lost following DNase treatment, they are thus identified as DNA molecules.

To demonstrate this, the poly(A)+RNA fraction was treated with alkali, under conditions which hydrolysed RNA, followed by glyoxalation and gel electrophoresis. Alkali-resistant single-stranded DNA bands up to 8Kb in size were detected (Fig. 2d). The pattern of small molecules was changed by this treatment. The abundant species sa- (725 nucleotides) was lost but the band corresponding to sb- (625 nucleotides) increased in intensity concomitantly (Fig. 2d). A similar result was obtained when the poly(A)+RNA fraction was treated with pancreatic RNase A before denaturing gel electrophoresis (Fig. 2e). These findings strongly suggest that sa- (designated sa-DNA) is a composite molecule containing covalently-linked ribonucleotides and that a

fragment of about 100 nucleotides was removed following each treatment. An additional small RNA fragment, which migrated much faster than the 0.2Kb size marker, was detected after DNase treatment of the poly(A)+RNA fraction (Fig. 2g, arrowed r). This is evidence that sa-DNA contains an uninterrupted stretch of about 100 ribonucleotides located at one end of the molecule.

Association of s-DNAs with other CaMV sequences. 1-D gel analysis. Because the CaMV small DNAs, and in particular the RNA-containing sa-DNA, co-purified with poly(A)+RNA, we wished to determine whether they were specifically associated with other nucleic acid species in this fraction. No small molecules were found as discrete bands in the poly(A)+RNA fraction blotted from a neutral gel (Fig. 3b). It appears, therefore, that the s-DNAs were sequestered by larger molecules. This association could be by duplex formation with either RNA or with other DNA species which we find co-purify with poly(A)+RNA. Digestion of the poly(A)+RNA fraction with RNase A released a small molecule (presumably DNA, because we find that RNA does not bind to nitrocellulose when blotted from a neutral gel), sometimes resolved into two closely-migrating components, from sequestration by larger molecules as revealed on a neutral gel (Fig. 3c, band arrowed). This result would be expected if s-DNAs were hybridised to larger RNA species. Since RNase A digestion removed the RNA component of sa-DNA, as shown by its increased mobility after denaturation (Fig. 2e), an alternative explanation

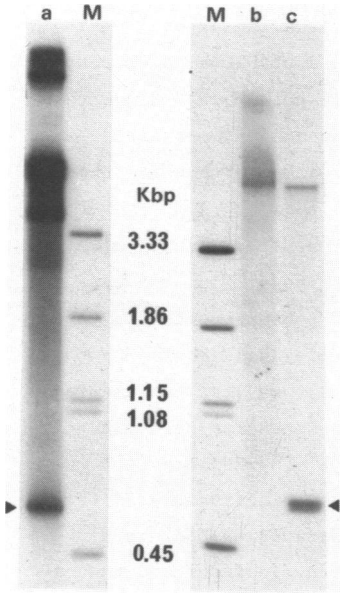

Figure 3. Non-denaturing gel electrophoresis and detection of CaMV-specific sequences, by blot-hybridisation,in (a), the cellular DNA fraction and (b), poly(A)+RNA treated with (c), RNase A. M, size markers. The arrowed bands are discussed in the text.

that sa-DNA is hybridised to other DNA molecules by its RNA moiety alone would seem unlikely; the latter association would be resistant to RNase A digestion in high salt because of the duplex structure.

When the salt-soluble (DNA) fraction was analysed on a non-denaturing gel, in addition to several high molecular-weight species, a CaMV-specific DNA molecule of 675 base pairs was detected (Fig. 3a, band arrowed).

2-D gel analysis. In two-dimensional gel electrophoresis, double-stranded nucleic acid species are first separated in a neutral buffer and then, at 90° to the first dimension, electrophoresis proceeds in an alkaline buffer which denatures double-stranded molecules (and hydrolyses RNA) and the single-stranded DNAs which arise from the double-stranded structures are sized relative to markers.

A small nucleic acid component, of 675 base pairs, from the salt-soluble (DNA) fraction was resolved by a neutral gel (see Fig. 3a) and was subsequently denatured to produce three DNA species of 625, 575 and 200 nucleotides aligned vertically in the alkaline dimension of a 2-D gel (Fig. 4a,b). These fragments correspond to the CaMV small DNAs sb-, sc- and sd-.

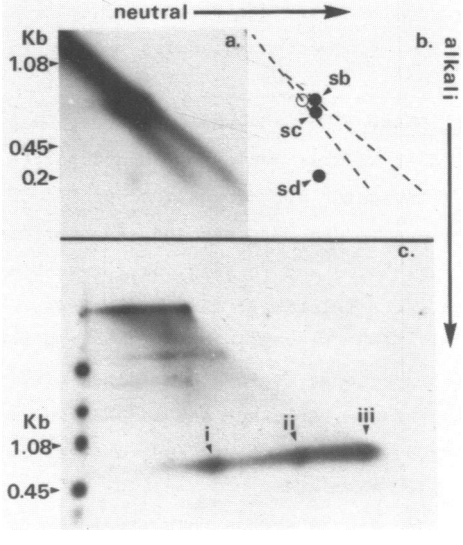

Figure 4. 2-D gel analysis of (a), the 675 base-pair CaMV-specific DNA molecule from the cellular DNA fraction. The film was overexposed to show sd-DNA. The interpretation of (a) is shown in (b). The open circle represents sa-DNA. (c), 2-D gel of the poly(A)+RNA fraction showing a horizonal tie-line of 0.62Kb produced by sa-DNA. The spots i, ii, iii are discussed in the text. Size markers in the alkaline dimension are to the left.

Thus, the small DNA found in the salt-soluble fraction appears to be a double-stranded molecule with three single-stranded components. The single-stranded components apparently lack ribonucleotides because their mobilities following denaturation by both glyoxal and alkali were similar. We have not detected any association of these molecules with higher molecular-weight species in the cellular DNA fraction.

The 2-D gel analysis of molecules present in the poly(A)+RNA fraction is shown in Fig. 4c. The long horizontal band, which is a 625 nucleotide DNA species sized in the alkali dimension relative to the markers shown to the left of Fig. 4c, represents sa-DNA, which has lost the RNA moiety due to hydrolysis in the alkali buffer. The interpretation of this result is that sa-DNA (the most abundant of the small DNAs in the poly(A)+RNA fraction) was hybridised to molecules of heterogenous size in the neutral buffer dimension and then released in the alkali buffer to migrate as a single molecular species of 625 nucleotides. In addition to its association with heterogenous molecules, 2-D gel electrophoresis further revealed that sa-DNA was hybridised to three molecules of discrete size (Fig. 4c, spots i, ii and iii). This association was only observed on a 2D gel because the second dimension effectively separated the discrete sized species from the heterogenous background of fragmented molecules present in the first (neutral) dimension.

Mapping the CaMV small DNAs. The CaMV small DNAs were mapped by blot-hybridisation of salt-soluble nucleic acid and poly(A)+RNA fractions, separated on glyoxal gels, using nick-translated probes of cloned CaMV Bgl II fragments. The single-stranded DNAs sa- and sb- hybridised strongly with Bgl II fragments -c (Fig. 5g,h) and -e (Fig. 5k,l). sc-DNA also hybridised with Bgl II-e (Fig. 5k,l). A relatively faint signal was obtained for sc-DNA when hybridised with Bgl II-c (Fig. 5g,h) because a low specific activity probe was used in this experiment. sd-DNA hybridised with Bgl II-e (Fig. 5k,l). No hybridisation of sa-, sb-, sc- and sd-DNAs was detected with Bgl II fragments -a (Fig 5c,d), -b (Fig. 5e,f) and -d (Fig. 5i,j) or with sd-DNA and Bgl II-c (Fig. 5g,h).

Attempts were made to map the s-DNAs more precisely using restriction enzymes. However, sa-DNA was refractory to digestion by all restriction enzymes tried. This finding reinforces the suggestion that sa-DNA is either single-stranded or, more likely, hybridised to RNA and therefore is not a substrate for restriction enzymes which recognise double-stranded DNA.

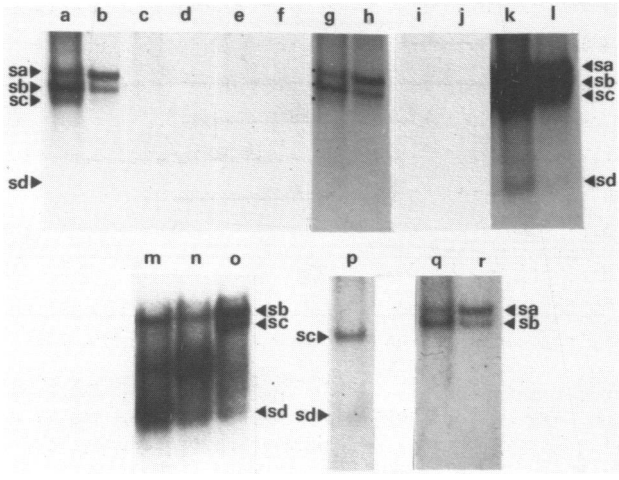

Figure 5. Mapping the small DNAs in the cellular DNA fraction (a, c, e, g, i, k, p, q) and poly(A)+RNA fractions (b, d, f, h, j, l, r,) by probing with total CaMV DNA (a,b), CaMV Bgl II fragments -a (c,d), -b (e,f), -c (g,h), -d (i,j) and -e (k,l) and virion DNA α-strand (p,) and β-strand (q,r). DNA in a sucrose gradient fraction containing (o), the 675 base-pair double-stranded DNA digested with (m), Hinf 1 and (n), Bgl II.

The small double-stranded DNA molecule of 675 base pairs (Fig. 3a), which was resolved by 2-D gel electrophoresis (Fig. 4a,b) into the three single-stranded components sb-, sc- and sd-, was purified by neutral sucrose gradient ultracentrifugation (Fig. 5o) for restriction enzyme analysis. Digestion of the double-stranded DNA in the gradient fraction with Bgl II produced a major band corresponding to a fragment of 360 nucleotides and the concomitant removal of sc-DNA and a reduction in the intensity of sb-DNA (Fig. 5n). Hinf 1 digestion produced a major fragment of 220 nucleotides and a minor one of 360 nucleotides (Fig. 5m). Blot-hybridisation suggested that sb- and sc-DNAs were complementary to the contiguous Bgl II fragments -c and -e (Fig. 5g,h,k,l); this conclusion is confirmed by the restriction enzyme mapping. The Bgl II-c/e junction is located at 0.955 m.u. on the CaMV genome and a Hinf 1 site is at 0.971 m.u.; therefore, common ends for sb- and sc-DNAs would be positioned at Gl in virion DNA.

The location of the other ends of these single-stranded molecules is not readily apparent as the fragments which should have been produced by enzyme digestion were not observed as major bands. This might indicate some heterogenity at one end of each molecule. Assuming common ends mapping at Gl, then by subtraction, the other ends of sb- and sc-DNAs are located at

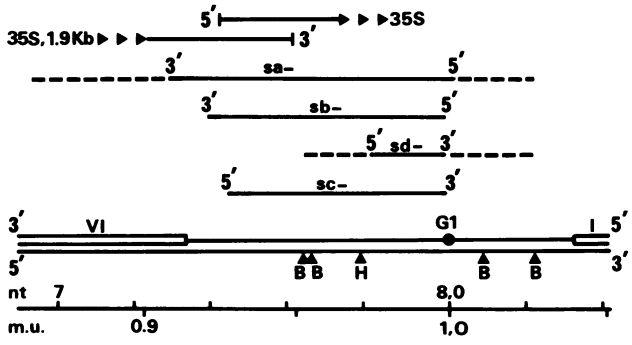

Figure 6. Alignment and polarity of the small single-stranded DNAs with that part of the CaMV genome from which they originate. The broken lines represent the limits of accuracy of the mapping procedures described in the text. IR1, which contains the DNA α-strand discontinuity G1, the 5'-terminus of 35S RNA and the 3'-co-termini of 35S RNA and the 1.9Kb mRNA, lies between open regions VI and I. The open regions have been drawn on the transcribed α-strand to indicate that they would appear in the RNA sense. Restriction enzyme sites: B, Bgl II; H, Hinf 1.

approximately 0.922 m.u. and 0.928 m.u. respectively. The sb-DNA component was not completely susceptible to restriction enzyme digestion (Fig. 5m,n) and this suggests that it may also exist in a single stranded form.

Hybridisation of sd-DNA was detected only with Bgl II-e (Fig. 5k,l) and since Bgl II digestion of the gradient fraction containing sd-DNA left this fragment uncut (Fig. 5n), the origin of sd-DNA is exclusively within Bgl II-e.

To determine the polarity of s-DNAs with respect to that of virion DNA, the separated α- and β-strands, labelled by nick-translation, were used as hybridisation probes. Both sc- and sd-DNAs hybridised with the α-strand (Fig. 5p) and are thus the same polarity as the β-strand. However, sa-and sb-DNAs hybridised with the β-strand (Fig. q,r) and therefore have the same polarity as the α-strand and opposite polarity to RNAs transcribed from the α-strand.

Figure 6 shows the location of the small DNA species, deduced from the mapping experiments described, aligned with that part of the CaMV genome from which they originate.

DISCUSSION

We have detected and partially characterised small DNA molecules which originate from the large intergenic region of the cauliflower mosaic virus genome.

Three small single-stranded DNAs (sb-, sc- and sd-), which co-purify with cellular DNA, appear to be associated with each other to produce a double-stranded molecule which migrates as a single band on non-denaturing gels (Fig. 3a; Fig. 4a,b). The single-stranded component sb-DNA (625 nucleotides) is the same polarity as the CaMV virion DNA α-strand and of opposite polarity to sc-DNA (575 nucleotides) and sd-DNA (200 nucleotides). Within the resolution of the mapping procedures used, we conclude that the 5'-end of sb-DNA and the 3'-end of sc-DNA map close to the position of the α-strand discontinuity (G1) found in virion DNA (Fig. 6). The difference in size between sb-DNA and sc-DNA (ca. 50 nucleotides) is not sufficient to accommodate sd-DNA in a double-stranded structure without producing a single-strand overhang or a D-loop similar to that found in some mitochondrial DNAs (18). sd-DNA must map within 200 nucleotides of G1 since it hybridised exclusively with the small Bgl II-e fragment which contains G1. This infers that sc-DNA and sd-DNA possess common sequences.

A further small single-stranded DNA, sa-DNA (725 nucleotides), co-purified with cellular RNA. It was shortened by 100 nucleotides following treatment with pancreatic RNase A and alkali and from this we conclude that sa-DNA contains covalently-linked ribonucleotides. An RNA fragment of approximately 100 nucleotides was observed following DNase treatment of the poly(A)+RNA fraction which suggests that the RNA moiety is located at one terminus of sa-DNA. This finding is evidence against the possible existence of a few ribonucleotides situated 100 nucleotides in from one end of the molecule. We do not yet know whether the RNA moiety is situated at the 3'-end or 5'-end of sa-DNA.

The polarity of sa-DNA was determined by hybridisation using strand-separated CaMV virion DNA probes for which the polarity is known (10). This showed that it is complementary to the β-strand and the RNA moiety must be produced, therefore, using the β-strand as template. This is the first evidence that the β-strand is transcribed into RNA, assuming that the RNA moiety is not of host origin. The major CaMV RNAs are transcribed from the α-strand (1,10) and since sa-DNA has the same polarity as the α-strand, it is thus of opposite polarity to the CaMV polyadenylated transcripts and could hybridise with them.

Results were presented which, in fact, showed that sa-DNA was preferentially associated with the cellular poly(A)+RNA fraction. It is possible that sa-DNA exists in the cell as a single-stranded molecule and during extraction became hybridised to complementary CaMV transcripts or

aggregated with cellular poly(A)+RNAs. Another possibility is that sa-DNA co-purified with other DNA molecules present in the poly(A)+RNA fraction. We do not know whether these other DNAs contaminate the poly(A)+RNA fraction or are more specifically associated with RNA.

Two-dimensional gel electrophoresis showed that sa-DNA was associated with molecules of heterogenous size in the poly(A)+RNA fraction, some of which had very low mobility in the neutral gel buffer. These molecules are, as yet, unidentified although we do have evidence that sa-DNA is specifically associated with molecules in the 35S region of a neutral sucrose gradient (data not shown). A discrete-sized molecule with relatively low mobility in the neutral dimension was associated with sa-DNA, as resolved on a 2-D gel (Fig. 4c, spot i), although we do not know whether it is related to 35S RNA because the mobility of 35S RNA in neutral gels has not been determined. Since the three spots (Fig. 4c; i, ii, iii) which fall along the major horizontal tie-line (corresponding to sa-DNA) produced in the alkaline dimension are not aligned below spots of equivalent intensity which would correspond to larger DNA species, we conclude that the molecules of discrete size, with which sa-DNA was associated, are RNAs.

The small DNAs found predominantly in the cellular DNA fraction were also present, but to a lesser extent, in the poly(A)+RNA fraction. This may have been due to some specific association with RNA or with sa-DNA.

We were unable to precisely map sa-DNA using restriction enzymes presumably because it does not exist as double-stranded DNA. But since it hybridised strongly to the adjacent Bgl II fragments -c and -e, then its 5'-terminus must lie fairly close to G1 and this would place its 3'-terminus at least as far upstream as the region transcribed to produce the 5'-terminus of 35S RNA (Fig. 6). As sa-DNA is complementary to 35S RNA and could hybridise with it, then two hybrid structures are possible because of the overlapping termini of 35S RNA. In one, sa-DNA might form a duplex exclusively with the 5'-end of 35S RNA. Alternatively, sa-DNA could hybridise with both ends of 35S RNA thus circularising it. Such structures might function by effecting some control over 35S RNA transcription, processing or expression. A more speculative possibility is that the circular molecule might be a replication intermediate with sa-DNA a primer hybridised to a 35S RNA template. If the products were RNA molecules then they would be complementary to the DNA β-strand. No such RNAs have been observed (with the possible exception of the RNA moiety of sa-DNA itself). In another scheme a DNA copy might be produced. However, we have no

results which demonstrate reverse transcription of 35S RNA although we have some evidence which suggests the existence of a single-stranded DNA with opposite polarity to the major CaMV transcripts (unpublished results).

Recently, Guilley et al.(19) have described two further CaMV RNAs with termini located in IR1. One is an 8S molecule with the same polarity as 35S RNA, which maps from the 5'-end of 35S RNA to G1 on the CaMV genome. This small RNA might be the 5'-end of 35S RNA, hybridised with sa-DNA, which has become detached by breakage or processing. A truncated form of 35S RNA would result with a 5'-end located at G1. Such a molecule has been observed as a minor component although its 3'-end also mapped at G1 (19). If a supercoiled form of CaMV DNA (11, 12) is transcribed continuously in a manner similar to that observed for polyoma virus (20) then sa-DNA might function by selecting fixed positions for processing giant transcripts. Our evidence from 2-D gel electrophoresis is that sa-DNA associates with heterogenous molecules, some with very low mobility in a neutral buffer (Fig. 4c). These could be tandemly repeated transcripts the existence of which has been inferred from our S1-mapping experiments (5).

The small DNAs which co-purify with cellular DNA are probably related to sa-DNA. Since they map close to G1, it is possible that they function in DNA replication. We have detected very low levels of the small DNAs (but not sa-DNA) encapsidated in CaMV virions (unpublished results). Whether they are primers for DNA replication, consequences of replication or function in some other process remains to be determined. The small DNAs may be related to other CaMV DNA forms that we have detected (Hull and Covey, in preparation) which appear to exist in a free or partially encapsidated state in infected cells.

ACKNOWLEDGEMENTS

We thank R Hull, J W Davies and A O Jackson for advice. G Mulder was a visiting student from the University of Wageningen, The Netherlands.

REFERENCES

1. Franck, A., Guilley,. H., Jonard, G., Richards, K. and Hirth, L. (1980). Cell 21, 285-294.
2. Gardner, R.C., Howarth, A.J., Hahn, P., Brown-Leudi, M., Shepherd, R.J. and Messing, J. (1981). Nucleic Acids Res. 9, 2871-2888.
3. Hohn, T., Richards, K. and Lebeurier, G. (1981). Current Topics in Microbiology and Immunology 96, 193-236.
4. Howell, S.H. (1982). Ann. Rev. Plant Phisiol. 33, 609-650.
5. Covey, S.N., Lomonossoff, G.P. and Hull, R. (1981). Nucleic Acids Res. 9, 6735-6747.

6. Dudley, R.K., Odell, J.T. and Howell, S.H. (1982). Virology 117, 19-28.
7. Covey, S.N. and Hull, R. (1981). Virology 111, 463-474.
8. Volovitch, M., Drudgeon, G. and Yot, P. (1978). Nucleic Acids Res. 5, 2913-2925.
9. Howell, S.H. and Hull, R. (1978). Virology 86, 468-481.
10. Hull, R., Covey, S.N., Stanley, J. and Davies, J.W. (1979). Nucleic Acids Res., 7, 669-677.
11. Menissier, J., Lebeurier, G. and Hirth, L. (1982). Virology 117, 322-328.
12. Olszewski, N., Hagen, G. and Guilfoyle, T.J. (1982). Cell 29, 395-402.
13. Hull, R. (1980). Virology 100, 76-90.
14. Covey, S.N. and Grierson, D. (1976). Europ. J. Biochem. 63, 599-606.
15. Favaloro, J.M., Triesman, R.H. and Kamen, R.I.(1980). Methods Enzymol. 65, 718-749
16. Southern, E.M. (1975). J. Mol. Biol. 98, 503-517.
17. Thomas, P.S. (1980). Proc. Nat. Acad. Sci. USA 77, 5201-5205.
18. Clayton, D.A. (1982). Cell 28, 693-705.
19. Guilley, H., Dudley, R.K., Jonard, G., Balazs, E. and Richards, K.E. (1982). Cell, In Press.
20. Triesman, R. and Kamen, R.I. (1981). J. Mol. Biol. 148, 273-301.