The Double-stranded DNA of Cauliflower Mosaic Virus

By G. J. RUSSELL, E. A. C. FOLLETT AND J. H. SUBAK-SHARPE
Institute of Virology, University of Glasgow, Church Street, Glasgow, W 1

AND B. D. HARRISON
Scottish Horticultural Research Institute, Invergowrie, Dundee

(Accepted 13 January 1970)

SUMMARY
Examination of the physical structure of cauliflower mosaic virus DNA by electron microscopy and ultracentrifugal analysis showed that the DNA could be isolated mainly as a circular molecule of mol. wt. $4.7 \times 10^6$. Nearest-neighbour frequency analysis showed that the DNA has a chemical structure which exhibits close similarities to that of the host cauliflower DNA. The relationships of this plant virus DNA and the small mammalian virus DNAs to the DNA of their respective hosts are discussed.

INTRODUCTION
Cauliflower mosaic virus (CIMV) and dahlia mosaic virus, which are serologically related, have a combination of attributes that distinguishes them from all other adequately studied plant viruses (Brunt, 1966). They have isometric particles about 50 nm. in diameter (Brunt 1966; Day & Venables, 1960; Pirone, Pound & Shepherd, 1960), they induce characteristic cytoplasmic inclusion bodies which appear amorphous in the light microscope but contain virus particles (Kitajima, Oliveira & Costa, 1965; Petzold, 1968), and they are transmitted by aphid vectors, in which they persist for only a few hours (Van Hoof, 1954; Brierley & Smith, 1950). In 1968, Shepherd, Wakeman & Romanko published evidence that CIMV contains DNA, the first plant virus shown to do so. Infective preparations of nucleic acid obtained from purified virus were found to contain thymine, and infectivity was abolished by treatment with deoxyribonuclease but not with ribonuclease. The cytoplasmic inclusion bodies in infected cells were subsequently shown to incorporate tritiated thymidine (Kamei, Rubio-Huertos & Matsui, 1969). In a further report Shepherd, Bruening & Wakeman (1970) concluded that the CIMV nucleic acid is double-stranded DNA, on the basis of its nucleotide composition, reactivity with formaldehyde and change of ultraviolet absorbency during heating.

Our work had two main aims: (1) to find out whether the nucleotide doublet composition of the virus nucleic acid is consistent with a two-stranded structure, and (2) to test the hypothesis of Subak-Sharpe et al. (1966) that viruses with small genomes have both nucleotide doublet patterns and general designs (Subak-Sharpe, 1967) that closely resemble those of the DNA of their hosts. The nucleic acids of large viruses have doublet patterns and general designs that may differ very considerably from host DNA. This hypothesis is compatible with the results obtained with several animal viruses and bacteriophages (Subak-Sharpe, 1969a, b), but has not previously been tested with any plant virus. In addition to providing answers to these questions, our results indicate that DNA molecules obtained from preparations of CIMV can have either a circular or a linear form.
METHODS

Preparation of virus and virus nucleic acid. To purify the virus, systemically infected leaves of tendergreen mustard (*Brassica perviridis*) were first disintegrated in pH 7.5 buffer (0.5 M-phosphate + 0.1 M-sodium diethyl dithiocarbamate; 1.5 ml./g. leaf tissue), using a mechanical blender. The extract was clarified by adding n-butanol to 8.5% (v/v), and the virus obtained by differential centrifugation, essentially as described by Pirone, Pound & Shepherd (1961). The partially purified virus preparations, suspended in 0.05 M-citrate buffer (pH 7-8), were then fractionated by exclusion chromatography using columns of agarose beads. Eluate from the middle of the virus-containing peak was collected and the virus concentrated by ultracentrifugation. The purified preparations had A260/A280 values of 1.41 to 1.42 and showed only one component, with a sedimentation coefficient of about 200 S, in the analytical ultracentrifuge. Electron microscopy showed that the preparations contained many 50 nm diameter virus particles and very little else. The yield of virus was usually 1 to 4 mg./100 g. plant issue.

DNA was prepared by treating the virus preparations with pronase and sodium dodecyl sulphate (R. J. Shepherd, personal communication). Pelleted virus particles were resuspended in 0.015 M-sodium chloride + 0.005 M-sodium citrate (= 0.1 SSC solution) at 2.5 mg. virus/ml. and pronase (B grade; obtained from Calbiochem Ltd.) was added to a concentration of 100 µg./ml. from a stock solution containing 10 mg./ml. (which had been pre-incubated for 3 hr at 37°). After incubation for 2 hr at 37°, the preparation had clarified considerably and no additional clarification was observed after further incubation for 1 hr. The preparation was then made 0.5 ~ with respect to sodium dodecyl sulphate and allowed to stand overnight at 20°.

In the initial bulk preparation, DNA was recovered from the pronase lysate by a modification of the Marmur (1961) procedure followed by isolation from CsCl density gradients. Subsequent DNA preparations were isolated from the pronase lysate directly on CsCl gradients.

The DNA was purified by banding twice on CsCl gradients. The DNA solution was diluted to 20 to 30 µg./ml. with 0.1 SSC, and solid CsCl (Harshaw Chemical Co., Optical grade) added to bring the density to 1.7 g./cm³. Solution densities were determined from refractive index measurements (Schildkraut, Marmur & Doty, 1962). The solution was then centrifuged at 40,000 rev./min. in a Beckman Ti 50 angle rotor for 60 to 70 hr at 20°. The density gradients were fractionated into 0.15 ml. fractions by piercing the tubes and collecting drops from the bottom. Fractions were made up to 1.0 ml. with 0.1 SSC and the DNA band located by measuring the absorbency of each fraction at 260 nm. Those fractions containing the DNA were pooled and, in the bulk preparation, the CsCl gradient procedure repeated. The final fractions containing the DNA were pooled and dialysed extensively against the 0.05 M-tris + 0.005 M-EDTA, pH 7.5. The final yield of DNA was about 72 µg./mg. virus.

Ultracentrifugal analysis of cauliflower mosaic virus DNA. Ultracentrifugal analyses were carried out on a Beckman Model E analytical ultracentrifuge using u.v. absorption optics.

Buoyant density determinations in CsCl were made as described by Schildkraut et al. (1962) using 0.5 to 1.0 µg. DNA + 0.5 µg. *Micrococcus luteus* DNA as marker (p = 1.731 g./cm.³) (Szybalski, 1968). CIMV DNA showed a single band of density 1.6696 g./cm.³ corresponding to a G+C content of 40.6%.

Sedimentation velocity experiments were done both by the moving boundary method
(Schachman, 1957), and the band sedimentation method (Vinograd & Brunner, 1966). In the moving boundary method, concentrations of 20 to 40 μg. of DNA/ml. of 0·15 M-NaCl were used. For band sedimentation, 0·5 to 2·0 μg. DNA in 0·01 to 0·05 ml. of low salt solution were layered on to either 1·0 M-NaCl or 3·0 M-CsCl. Band sedimentation in CsCl was used only to observe heterogeneity and for comparison between DNA preparations. Sedimentation coefficients were determined from moving boundary analyses and from band sedimentation through 1·0 M-NaCl.

Electron microscopy of cauliflower mosaic virus DNA. The purified CIMV DNA was prepared for examination by the protein monolayer technique (Kleinschmidt & Zahn, 1959). A sample containing approximately 1 μg. was mixed with 0·9 ml. m-ammonium acetate and 0·1 ml. cytochrome c (concentration 1 mg./ml. in m-ammonium acetate) and 0·4 ml. of this mixture allowed to spread slowly over the surface of a solution of 0·15 m-ammonium acetate. Areas of the protein film were picked up, dried in ethanol + isopentane and shadowed with Pt/Pd at an angle of 5° while rotating.

Preparation of DNA and cauliflower plants. DNA was prepared from fresh cauliflower curd by a modification of the method of Lytleton & Petersen (1964). The curd (~150 g.) was chopped up and suspended in 500 ml. of 0·1 M-NaCl + 0·025 M-tris + 0·01 M-EDTA, pH 7·5. This suspension was homogenized in a Servall Omnimix for 30 sec. using the 400 ml. capacity stainless steel vessel at variac setting 180. The resultant suspension was filtered through Whatman no. 1 paper and the retained material resuspended in 150 ml. 1% sodium dodecyl sulphate in 0·025 M-tris, pH 7·3, and this suspension stirred with an equal volume of 90% (w/v) phenol at room temperature for 60 min. The layers were separated by centrifugation at 1000g for 10 min. Two volumes of ethanol were added to the aqueous phase and the precipitate allowed to settle for 60 min at 4° and then collected by centrifugation at 1000g for 15 min.

The precipitate was redissolved in 20 ml. SSC and an equal volume of 4 M-sodium chloride added. This solution was left overnight at 4° and then clarified by centrifugation at 1000g for 15 min. Ethanol (1·5 vol.) was added to the supernatant fluid and the resultant precipitate collected by centrifugation at 1000g for 10 min. This precipitate was redissolved in 10 ml. 0·1 SSC and the DNA isolated from the solution by banding on CsCl density gradients as described previously. The final pooled fractions of DNA were dialysed extensively against 0·05 M-tris + 0·005 M-EDTA, pH 7·5. The yield of DNA was about 1 mg./g. of cauliflower curd. On buoyant density centrifugation in CsCl this DNA showed a single band of density 1·695 g./cm.³, corresponding to a G+C content of 36%.

Preparation of DNA from baby hamster kidney cells. DNA was isolated from monolayers of BHK 21/C13 cells by the method of Marmur (1961). On buoyant density centrifugation in CsCl this DNA showed a density of 1·705 g./cm.³, corresponding to a G+C content of 42%.

Nearest neighbour analysis of base sequence in DNAs. The CIMV DNA, cauliflower DNA and BHK 21/C13 DNA were all subjected to nearest-neighbour frequency analysis using the method of Josse, Kaiser & Kornberg (1961), modified as described by Subak-Sharpe et al. (1966). Other minor modifications include: (1) addition of sodium arsenate to a concentration of 2·5 mM in place of phosphate or NaF, as a phosphomonoesterase inhibitor in the spleen phosphodiesterase digestion (McGeoch, Crawford & Follett, 1970); and (2) separation of the deoxyribonucleoside-3'-monophosphates by descending chromatography on Whatman no. 1 paper developed in isobutyric acid: water: NH₄OH (sp.gr.0.880) (66:33:2) for 24 to 30 hr. This solvent gives a slightly better resolution between 3'-dGMP and 3'-dTMP than the less alkaline solvent (66:33:1) normally used.
The location of the separated 3'-deoxyribonucleotides was determined by examination under u.v. light, and the radioactivity estimated by immersing each spot in 10 ml. toluene-based scintillation fluid and counting in a scintillation counter. The remainder of the chromatogram was also counted in sections to determine the possible presence of any DNA residues not completely reduced to mononucleotides or the presence of radioactive inorganic phosphate indicating the action of phosphomonoesterase.

**RESULTS**

*Physical characterization of CIMV DNA*

(a) *Ultracentrifugal Analysis.* Two components of 18s and 20s were observed. The relative amounts of the 18s and 20s components varied in different preparations. In most earlier preparations they were in roughly 1:1 proportions. In the later preparations, which were isolated directly on CsCl gradients, more than 90% of the DNA was in form of the 20s component. However, after such preparations were stored at −20°C for a week, all the DNA was converted into the 18s component. On prolonged storage of all the DNA preparations...
at -20°, the sedimentation pattern changed to form a single broad boundary of much decreased s value.

Sedimentation velocity experiments on the 20s component run at pH 12 to 13 showed a somewhat diffuse single boundary sedimenting at 17·3s, corresponding to an average mol. wt of 1·93 x 10^6 (Studier, 1965). At no time was a faster moving component observed.

(b) Electron microscopy. Samples which by centrifugal analysis showed both 18 and 20s components contained both linear and circular molecules in varying proportion (Fig. 1). Examination of the later preparations with the 20s component in excess (90 %) showed that more than 80 % of the molecules were circular. No tightly supercoiled molecules were observed, such as are present in the component 1 fraction of preparations of polyoma virus and human papilloma virus DNA (Crawford, 1965). However, nearly all molecules of this C1MV DNA had one or more cross-over point and completely open circles with no cross-over points were exceptional. Measurement of 37 circular molecules with not more than three cross-over points gave a mean molecular length of 2·47 ± 0·14 μm. (Fig. 2), which corresponds to a mol. wt of 4·74 x 10^6, assuming a value of 1·92 x 10^6/μm. for the DNA. This value is very similar to the estimate of 4·6 x 10^6 calculated for the 18s component (Studier, 1965), assuming that this is the linear double-stranded form and that the 20s is the circular double-stranded form. Measurements of 43 linear molecules showed no molecule longer than the longest circular molecule and no obvious peak in the size distribution of the fragmented molecules (Fig. 2).

Nearest neighbour analysis of C1MV-DNA and cauliflower DNA

The results of the analyses of C1MV DNA, cauliflower DNA and BHK 21/C13 DNA are shown in Table 1, together with the analysis of wheatgerm DNA, published by Swartz, Trautner & Kornberg (1962), and that for polyoma virus DNA published by Subak-Sharpe et al. (1966). The base compositions determined from this data are also shown in Table 1. Those for C1MV DNA suggest a G+C content somewhat lower than that determined by buoyant density measurement.
Table 1. Doublet frequencies expressed as parts/1000

<table>
<thead>
<tr>
<th>Doublet</th>
<th>Wheatgerm*</th>
<th>Cauliflower mosaic virus</th>
<th>BHK 21/C13 Cells</th>
<th>Polyoma† virus</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApT TpA</td>
<td>75 58</td>
<td>104 88</td>
<td>86 74</td>
<td>79 71</td>
</tr>
<tr>
<td>ApA TpT</td>
<td>72 89</td>
<td>108 131</td>
<td>115 118</td>
<td>95 98</td>
</tr>
<tr>
<td>GpT ApC</td>
<td>56 53</td>
<td>52 42</td>
<td>40 45</td>
<td>56 55</td>
</tr>
<tr>
<td>TpG CpA</td>
<td>70 68</td>
<td>64 55</td>
<td>51 60</td>
<td>73 77</td>
</tr>
<tr>
<td>GpA TpC</td>
<td>62 71</td>
<td>59 58</td>
<td>61 77</td>
<td>58 65</td>
</tr>
<tr>
<td>ApG CpT</td>
<td>60 67</td>
<td>55 55</td>
<td>64 76</td>
<td>66 74</td>
</tr>
<tr>
<td>GpG CpC</td>
<td>51 59</td>
<td>42 36</td>
<td>35 46</td>
<td>44 50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Doublet</th>
<th>Frequency analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>GpC CpG</td>
<td>(G+C)%</td>
</tr>
<tr>
<td>a t</td>
<td>0.260 0.287</td>
</tr>
<tr>
<td>g c</td>
<td>0.220 0.233</td>
</tr>
<tr>
<td>a/t g/c</td>
<td>0.906 0.957</td>
</tr>
<tr>
<td>(G+C)%</td>
<td>45.3 35.1</td>
</tr>
<tr>
<td>Buoyant density</td>
<td>36.0 40.6 42.0 48.0</td>
</tr>
</tbody>
</table>

* Values for DNA taken from Swartz et al. (1962).
† Values for DNA taken from Subak-Sharpe et al. (1966).

Fig. 4. Histogram of superimposed doublet frequencies of cauliflower and CMV DNAs. Frequencies are expressed in parts/1000. Solid areas represent excess in virus over host. Stippled areas represent excess in host over virus.
Table 2. Doublet frequencies normalized to 50% G + C and expressed as parts/1000

<table>
<thead>
<tr>
<th></th>
<th>Wheatgerm*</th>
<th>Cauliflower</th>
<th>Cauliflower mosaic virus</th>
<th>BHK 21/ C 13 Cells</th>
<th>Polyoma virus</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApT TpA</td>
<td>63</td>
<td>61</td>
<td>54</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td>ApA TpT</td>
<td>67</td>
<td>52</td>
<td>75</td>
<td>66</td>
<td>66</td>
</tr>
<tr>
<td>GpT ApC</td>
<td>55</td>
<td>49</td>
<td>64</td>
<td>46</td>
<td>56</td>
</tr>
<tr>
<td>TpG CpA</td>
<td>56</td>
<td>54</td>
<td>72</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td>GpA TpC</td>
<td>55</td>
<td>70</td>
<td>66</td>
<td>49</td>
<td>52</td>
</tr>
<tr>
<td>ApG CpT</td>
<td>66</td>
<td>52</td>
<td>74</td>
<td>52</td>
<td>62</td>
</tr>
<tr>
<td>GpG CpC</td>
<td>65</td>
<td>80</td>
<td>70</td>
<td>66</td>
<td>64</td>
</tr>
</tbody>
</table>

* Values calculated from data of Swartz et al. (1962)

Fig. 5. General design histograms of wheatgerm, cauliflower and CIMV DNAs. Doublet frequencies are normalized to 50% G + C and expressed in terms of deviation, in parts/1000, from random expectation (62.5 parts of each doublet). G + C contents are given values calculated from nearest neighbour analyses.

The nearest neighbour frequency patterns can be compared conveniently by directly superimposing doublet frequency histograms as shown in Fig. 3 and 4. In this comparison no allowance is made for actual differences in the G + C content of the DNAs. The effect of differences in G + C content can be removed. To do this, the doublet frequencies are normalized according to the method of Subak-Sharpe et al. (1966), to the values they would have
if the DNAs had a G + C content of 50%. Table 2 gives the normalized frequencies for the DNAs listed in Table I. These results can be plotted as deviations from random expectation giving the 'general design' of the DNAs (Fig. 5, 6).

The sixteen doublet frequencies obtained in nearest neighbour frequency analysis of double-stranded DNAs, can be separated into two categories. These are (1) the four independent doublets ApT, TpA, GpC and CpG which are their own anti-parallel complement and (2) the dependent doublets which are complementary in the anti-parallel strands and should therefore be present with the same frequency; they are: ApA = TpT, GpT = ApC, TpG = CpA, GpA = TpC, ApG = CpT and GpG = CpC. As can be seen from Table I, the data for both cauliflower DNA and for CIMV DNA show reasonable agreement between their respective dependent doublets. When the frequencies are normalized to 50 % G + C (Table 2 and Fig. 5), the complementary dependent doublets show very good agreement in their pattern of deviation from random expectation, both in cauliflower DNA and in CIMV DNA. This confirms the conclusions of Shepherd et al. (1970) that CIMV DNA is double-stranded.

The results for the BHK 21/C13 DNA were essentially the same as previously reported (Subak-Sharpe et al. 1966), but improvement in the experimental technique has yielded data with better agreement between the complementary doublets (Fig. 6).

**DISCUSSION**

Our results show that, under the gentlest extraction conditions used, CIMV DNA exists as a circular molecule with a mol. wt of $4.7 \times 10^6$ sedimenting at 20s. This circular molecule is easily converted to a linear form sedimenting at 18s.
Human papilloma virus DNA has also been shown to exist in circular and linear forms sedimenting at identical values (Crawford, 1965). However, these forms have been shown to be derived from a supercoiled form which sediments at 28s. No evidence was obtained, in any of our experiments with CIMV DNA, of any component sedimenting faster than 20s. Supercoiled molecules are readily converted into circular molecules by a single scission in either strand. A circular molecule so derived would be expected to show two components while sedimenting in alkaline conditions (i.e. one circular and one linear strand) (Crawford, 1965). Our circular CIMV DNA, sedimenting in the presence of alkali, exhibits only one component of less than the expected half molecular weight. From this it seems probable that the CIMV DNA has several scissions in both strands.

This behaviour in alkali, together with the poor keeping qualities of the DNA, suggest that a nuclease may be associated with the DNA. If this is so, then supercoiled molecules could be present in the virus particle but would not remain intact during the extraction procedures.

Wheatgerm DNA has already been shown to have an appreciably different doublet frequency pattern from the established mammalian pattern (Swartz, et al. 1962). This has been confirmed by the analysis of cauliflower DNA in which the doublet frequency and the general design patterns show close similarities. As these two DNAs are from unrelated plants (monocotyledon and dicotyledon), it leads us to suggest that this pattern will prove representative of the higher plants. Unpublished results from a lower plant (Chlamydomonas) show a pattern with distinct differences.

The doublet frequency analysis of C1MV-DNA, as well as confirming the double-stranded nature of the DNA, shows that the general design of this DNA resembles that of the host cell. Subak-Sharpe et al. (1966) have already shown that the general design of the nucleic acid of small mammalian viruses resembles that of their host cells. They argue that these similarities occur because these small viruses have to make use of pre-existing host translation apparatus for polypeptide synthesis and probably originate from the DNA of the ancestral host cells. Our work shows that the general design of the DNA of one small plant virus resembles that of its host cell, but we think it premature to draw general conclusions from this single comparison. We must point out that in this plant system there is no highly characteristic doublet, such as the CpG shortage which exists in the mammalian system, that can be used as a reference. Care must therefore be exercised when making comparisons.

We wish to thank Dr R. J. Shepherd, University of California, Davis, for details of his methods of purifying the intact virus and virus nucleic acid, and Mr D. E. Branney and Mrs Margaret Scott for technical assistance.

REFERENCES


(Received 24 November 1970)