Molecular cloning and characterization of a complete DNA copy of potato spindle tuber viroid RNA

Peter van Wezenbeek*, Pieter Vos, Jacques van Boom⁺ and Albert van Kammen[§]

Department of Molecular Biology, Agricultural University, De Dreijen 11, 6703 BC Wageningen, The Netherlands

Received 28 October 1982; Accepted 16 November 1982

SUMMARY

Double-stranded cDNA has been synthesized from RNA of a severe strain of potato spindle tuber viroid using a synthetic oligodeoxyribonucleotide as a primer. Upon cloning in bacteriophage M13mp9, two recombinant phages were selected to construct a pBR322-derived plasmid containing a complete viroid DNA copy. Elucidation of the nucleotide sequence revealed four differences with the previously established sequence of another PSTV strain, three of which were base exchanges and one a deletion.

INTRODUCTION

Viroids are small single-stranded circular RNA molecules causing several economically important diseases in cultivated plants (1). The complete molecular structures of several viroids have been established including potato spindle tuber viroid (PSTV, ref 2), chrysanthemum stunt viroid (CSV, ref 3), citrus exocortis viroid (CEV, ref 4), avocado sunblotch viroid (ASBV, ref 5) and cadang-cadang viroid (6). These RNAs range in size from 246 to 371 residues and form extended rod-like structures with a high degree of intramolecular base-pairing.

In its most severe form potato spindle tuber viroid (PSTV) causes general stunting of potato plant growth, deformity of the upper foliage and production of disfigured potatoes (7). As potatoes are vegetatively propagated, it is of major importance to exclude viroids from seed potatoes. Currently, detection of PSTV involves laborious and expensive assays including polyacrylamide gel electrophoresis of extracted nucleic acids (8, 9). Very recently, however, Owens and Diener (10) described an alternative method which is based on hybridization of radioactively labeled PSTV cDNA with RNA extracts from clarified plant sap bound to a solid support. This method is attractive since it is simple, sensitive, inexpensive and reproducible. On request of the Dutch Plant Protection Service (PD) and the Research Institute for Plant Protection (IPO), we constructed a plasmid containing a full-length DNA copy of PSTV which may

serve as a hybridization probe in such a nucleic acid spot hybridization procedure.

Comparative sequence analyses of severe and mild PSTV strains have indicated that their nucleotide sequences are very similar and that only few nucleotide differences occur (11, 12). Generally, to be able to correlate nucleotide changes in a defined region with a particular symptom pattern in the plant, knowledge of structural details of a great number of isolates will be required. In this respect, it was also of interest to establish the nucleotide sequence of the PSTV strain used in this study which evokes the symptoms characteristic for a severe strain. Moreover, construction of a full-length PSTV DNA copy was of great importance since it might be successfully used to produce viroid molecules upon transformation of potato cell protoplasts. If indeed such a clone generates biologically active viroids analogous to what has been described for full-length DNA copies of phage Q β and poliovirus RNAs (13, 14), this would enable a thorough study on the nature of PSTV pathogenicity by site-directed mutagenesis on DNA level.

MATERIALS AND METHODS

Enzymes. *E. coli* DNA polymerase I was obtained from Boehringer, Mannheim. T4 DNA ligase was prepared from an induced lysogen of λ T4lig phage NM989 (15) and further isolated with minor modifications according to the method of Panet *et al.* (16). T4 polynucleotide kinase was purchased from P.L. Biochemicals and the restriction endonucleases used in this study were from Boehringer or Bethesda Research Laboratories Inc. AMV reverse transcriptase was obtained from J.W. Beard, Institute of Life Sciences, St. Petersburg.

Preparation of nucleic acids. The severe strain of PSTV used in this study (isolated from *Solanum commersonii* Dun. ref 17) was originally obtained from P. Howell, Edinburgh. PSTV RNA, a kind gift of W. Mosch, Wageningen, was purified from infected tomato tissue (*Lycopersicon esculentum* Mill., cv Rutgers, ref 18). Final purification involved two cycles of polyacrylamide gel electrophoresis, the second of which was performed under denaturing conditions to obtain the circular form of PSTV (19). To remove gel impurities, purified RNA was bound to DEAE-Sephacel (Pharmacia) in 50 mM Tris-HCl, pH 8.0, washed with the same buffer and eluted with buffer containing 1.2 M NaCl.

Plasmid DNA was isolated essentially as described by Godson and Vapnek (20). The plasmid DNA finally obtained was extracted twice with an equal volume of phenol, three times with diethylether and precipitated with ethanol. Recombinant M13 RF DNA was isolated from infected cells in a similar way.

The 15-mer d(TTCTTTTTTCTTTTC) and the 18-mer d(GTAAAACGACGGCCAGTG) were synthesized by the chemical phosphotriester method (21).

Synthesis of double-stranded cDNA. Single-stranded PSTV cDNA was synthesized in a reaction mixture (10 μ l) containing 50 mM Tris-HCl (pH 8.3), 80 mM KCl, 10 mM MgCl₂, 1 mM DTT, 0.05% NP-40, 0.5 μ g of circular PSTV, 1.5 μ g of primer, 0.5 mM each of dGTP, dCTP and dTTP, 0.1 mM labeled (α -³²P)dATP (specific activity 10 Ci/mmol) and 20 units of reverse transcriptase. Prior to DNA synthesis, template PSTV RNA was heated for 1 min at 100°C in the presence of the chemically synthesized primer d(TTCTTTTTTCTTTTC) and immediately quenched in ice-water. Enzymatic reactions were carried out at 37°C for 2 h. After heat denaturation, PSTV template RNA was degraded by addition of RNase A to the reaction mixture (final concentration 10 μ g/ml) and incubation for 30 min at 37°C. Single-stranded PSTV cDNA was purified on a Sephadex G-50 column and incubated with 5 units of DNA polymerase I in a volume of 50 μ l containing 100 mM potassium phosphate (pH 7.0), 6.5 mM MgCl₂, 1 mM DTT and 250 mM each of the four deoxyribonucleoside triphosphates. The reaction was carried out at 37°C for 2 h.

Construction of recombinant DNA molecules. Chemically synthesized DNA octamers containing an *Eco*RI endonuclease cleavage site (Collaborative Research) were phosphorylated at their 5'-end with ATP and T4 polynucleotide kinase by incubation for 45 min at 37°C in a 10 μ l mixture containing 25 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 10 mM DTT, 0.1 mM ATP, 20 pmoles of linker molecules and 5 units of enzyme. The reaction mixture was adjusted to 1 mM ATP and unfractionated cDNA and 1 μ l of T4 DNA ligase (the optimal amount of enzyme needed to ligate blunt-ended DNA fragments) were added. Blunt-end ligation was performed at 15°C by overnight incubation. Generation of cohesive ends was performed by digestion with 10 units of endonuclease during 1 h at 37°C. Subsequent ligation to 50 ng of linearized M13mp9 RF DNA was carried out in a reaction volume of 20 μ l containing 25 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 10 mM DTT and 0.2 mM ATP. Ligation was performed by overnight incubation at 15°C using an amount of T4 DNA ligase which was 10 times less than in the blunt-end ligation reaction. Construction of the complete PSTV cDNA-containing plasmid pAV401 was performed under blunt-end ligation conditions by adding 0.1 μ g of linearized pBR322 to a twenty-fold molar excess of fragments which were isolated from M13mp9 recombinant phages.

Competent *E. coli* HB101 or JM103 cells were used for transformation or transfection, respectively. Transformants arisen by uptake of plasmids were selected by plating on nutrient agar containing 20 μ g/ml ampicillin, whereas trans-

fected cells were plated in soft agar in the presence of 0.16 mg/ml isopropylthio-galactoside and 0.2 mg/ml 5-bromo-4-chloro-3-indolyl- β -D-galactoside to select for recombinants causing insertional inactivation of the production of β -galactosidase (22).

DNA sequencing. The PSTV sequence was determined by using the dideoxynucleoside triphosphate chain-termination method described by Sanger *et al.* (23). Limited DNA synthesis was performed on M13mp9 recombinant phage DNA (22, 24) annealed with the "universal" primer d(GTAAACGACGGCCAGTG). Standard procedures were used for preparing single-stranded phage DNA (25).

RESULTS AND DISCUSSION.

Synthesis of double-stranded PSTV cDNA. It has been demonstrated previously that the pentecaidecadeoxyribonucleotide d(TTCTTTTCTTTTC), which is complementary to the nucleotide numbers 49-63 of PSTV, can successfully be used as a primer for the synthesis of the first strand (26). The sizes of cDNA synthesized using this primer are shown in Fig. 1b. A distinct prominent band can be seen representing full-length cDNA. No single-stranded cDNA synthesis was detected in the absence of the primer.

The short hairpin loops present at the 3'-termini of single-stranded DNA were used to prime the synthesis of the second DNA strand by DNA polymerase I. The products of this synthesis were analyzed by electrophoresis on a 5% polyacrylamide gel under denaturing conditions (Fig. 1c). A number of discrete cDNA classes are visible, the slowest migrating band containing fragments of approximately 710 bases as would be expected for full-length unfolded double-stranded PSTV cDNA.

A further characterization was obtained by testing the sensitivity of the unfractionated reaction products to cleavage by different endonucleases. The largest fragments which should arise after cleavage with restriction enzymes *Bam*HI, *Hae*III or *Hha*I have lengths of approximately 330, 280 and 175 nucleotides, respectively, as deduced from the nucleotide sequence (ref. 2, Fig. 5). Indeed, fragments of the predicted lengths were found upon digestion of the double-stranded cDNA with these enzymes (Fig. 2). It should be noted that the unique *Bam*HI site is located at a distance of only 24 nucleotides from the 5'-end of the primer sequence. The slight difference in electrophoretic mobility between the most prominent bands in the patterns of *Bam*HI-treated (lane a) and untreated cDNA (lane d) again indicates that the products synthesized were almost of full-length size.

Molecular cloning of double-stranded PSTV cDNA. Since the *Bam*HI recognition

Fig. 1. Analysis of the reaction products obtained after PSTV cDNA synthesis. Aliquots of the reaction mixtures of the first strand synthesis (lane b) and the second strand synthesis (lane c) were layered in 80% formamide, 8 mM EDTA on a denaturing 6% polyacrylamide gel containing 7 M urea. Prior to loading, samples were heated for 2 min at 90°. Electrophoresis was performed at 250 V during 3 h using Tris-borate buffer (23). Lane a contains size markers which were obtained by digesting M13 RF DNA with endonuclease *Hae*III (27).

site is located at only a short distance from the terminal loop formed after the first strand synthesis we used this site for removing that loop by digestion with endonuclease *Bam*H1. This had the advantage over the conventionally used nuclease S1 treatment in that the linear double-stranded cDNA had one cohesive end which can be directly ligated into a suitable vector molecule. The other end of the cDNA was expected to be blunt.

Two approaches were followed for molecular cloning. Firstly, the *Bam*H1 digested cDNA was ligated to M13mp9 RF cleaved with *Bam*H1 and *Sma*I. Secondly, in order to obtain cohesive ends at both termini of the cDNA, *Eco*RI linker molecules were attached to the cDNA prior to *Bam*H1 digestion. Cohesive ends were subsequently generated by cleavage with *Eco*RI and *Bam*H1. These molecules were then inserted into M13mp9 RF linearized by the same two endonucleases. Both ligation mixtures were used to transform competent *E.coli* cells and white plaques were selected on nutrient agar plates. Ultimately, three recombinants were grown up for further investigation. They were called M13AV104 (isolated from the first cloning method) and M13AV305 and M13AV312 (derived from the second cloning method).

The advantage of using bacteriophage M13mp9 as a vector is that single-stranded recombinant phage DNA can easily be isolated and used as a template in the rapid dideoxy sequencing method (23, 25). Determination of the nucleo-

Fig. 2. Restriction enzyme analysis of double-stranded PSTV cDNA. Double-stranded cDNA (see Materials and Methods) was electrophoresed under nondenaturing conditions on a 6% polyacrylamide gel and visualized by autoradiography (lane d). Lanes a, b, c: double-stranded PSTV cDNA digested with endonucleases *Bam*H1, *Hae*III or *Hha*I, respectively.

tide sequence allowed us to define precisely the lengths and origins of the inserts. The insert of M13AV104 appeared to have a length of 290 nucleotides and originated from the end of the *Bam*H1 site distal to the priming site (Fig. 3). In M13AV305 a sequence of 334 nucleotides of PSTV cDNA was found to be integrated which was the sequence extending from the 5'-end of the primer towards the *Bam*H1 site in the 5' → 3' direction of the complementary strand. It was expected that this recombinant phage would contain the maximum length of PSTV sequence possible if *Bam*H1 is used to remove the terminal hair-pin. Thus, the 24-nucleotide fragment located between the *Bam*H1 site and the primer sequence could not be cloned.

However, M13AV312 contained 351 PSTV-characteristic nucleotides originating from the primer-proximal end of the *Bam*H1 site and including therefore this 24-nucleotide fragment. The existence of this recombinant can only be explained by assuming that occasionally the enzyme reverse transcriptase does not stop after one round of synthesis but, instead, proceeds through the primer region thereby displacing the original primer-extended first strand. Digestion of this longer-than-unit-length double-stranded cDNA will then result in two fragments, one of which extending from the primer sequence to the unique *Bam*H1 site and identical to the insert of M13AV305. The second fragment ex-

Fig. 3. Circular representation of the covalently closed circular RNA sequence of PSTV. In the inner circle several restriction endonuclease cleavage sites present in double-stranded cDNA are marked. The position where the primer hybridizes to the RNA and the direction of the first strand synthesis are indicated by an arrow. The outer circles denote the regions integrated in the recombinant phages indicated. The rearranged sequences contained in M13AV312 which are located below position 181 of the PSTV sequence (see text) are not outlined.

tends from the *Bam*HI site through the primer region to the position where the hairpin structure is formed to prime the synthesis of the second strand. Second strand synthesis can be followed by nicking of the loop and filling in the single-stranded end by DNA polymerase I (28). Insertion of such a fragment in M13mp9 can account for the recombinant bacteriophage M13AV312. According to the mechanism proposed by Volckaert *et al.* (28), opening the loop by some nicking activity would lead to DNA rearrangements. In fact, rearrangements have been found in the insert of M13AV312 which are located next to the *Eco*RI site (in total 76 nucleotides).

As we were interested in constructing a recombinant DNA molecule with a complete PSTV cDNA insert we isolated replicative form DNA of two recombinant phages (M13AV312 and M13AV104) which together contained the whole PSTV sequence. The *Hind*III-*Alu*I fragment of M13AV312 RF DNA and the *Alu*I-*Bam*HI fragment of M13AV104 RF DNA were purified and ligated together in the presence of plasmid pBR322 which was first digested with endonucleases *Bam*HI and *Hind*III (Fig. 4). The mixture was used to transform competent *E. coli* cells. As the tetracycline resistance gene of pBR322 was destroyed by the exchange of the *Bam*HI-*Hind*III fragment whereas the ampicillin resistance gene remained unaffected, recombi-

Fig. 4. Schematic diagram of the construction of plasmid pAV401 bearing the complete PSTV cDNA sequence. The two fragments of M13AV104 and M13AV312 which were used for the construction of pAV401 are underlined. Open boxes indicate PSTV cDNA sequences. Double-lines represent sequences derived from M13mp9. The single line refers to the pBR322 sequence. The rearranged PSTV cDNA sequences present in M13AV312 are indicated by a hatched region. It should be mentioned that the orientation of the insert contained in M13AV104 is opposite to the theoretically expected one. Several restriction enzyme cleavage sites are denoted: H, *Hind*III; A, *Alu*I; B, *Bam*HI; E, *Eco*RI. P refers to the position of the primer sequence.

nant plasmids would only confer ampicillin resistance to transformed *E. coli* cells. Transformants could therefore be selected on basis of their resistance to ampicillin and sensitivity to tetracycline. One of the transformants contained a plasmid referred to as pAV401. Characterization of this plasmid revealed that the *Hind*III and *Bam*HI cohesive ends and the blunt *Alu*I ends used in the ligation procedure were joined correctly and that these restriction sites were restored. From alterations in the digestion patterns of *Alu*I, *Hha*I and *Hpa*II as compared to pBR322 it could be unambiguously concluded that plasmid pAV104 contained the complete PSTV cDNA sequence (data not shown). The insert could be excised precisely with endonuclease *Bam*HI.

Nucleotide sequence. The nucleotide sequence of the PSTV insert in plasmid pAV401 is presented in Fig. 5. It was deduced from the sequences of the M13 recombinant phages M13AV104, M13AV305 and M13AV312 which were determined by making use of the dideoxynucleoside triphosphate sequencing method (23). The inserts of these clones were completely sequenced and agreed with each other. Thus, cloning errors due to the reverse transcriptase reaction can readily be ruled out. Furthermore, with the exception of the 24 bases located between the primer sequence and the *Bam*HI site which were integrated only in M13AV312, the sequence was independently determined in both strands.

Compared to the nucleotide sequence of the severe PSTV strain described by Gross *et al.* (2), 4 nucleotide differences have been found in the *S. commerson-*

Fig. 5. Nucleotide sequence of the complete PSTV cDNA insert contained in plasmid pAV401 (located between the two *Bam*HI sites). Nucleotide changes in the severe strain determined by Gross *et al.* (2) are denoted on top of the line, differences in the mild strain sequence (12) are shown beneath each line. The PSTV sequence indicated is homologous to the viroid RNA whereas the pBR322 part is shown in an anti-clockwise direction (29). Due to the strategy followed for the construction of pAV401 a small piece (20 bases) of M13mp9 is conserved. The numbering of the nucleotides refer to the PSTV sequence only. The asterisk denotes the reference point. The underlined sequence denoted by P is complementary to the primer sequence. The sequences of several endonuclease recognition sites are indicated by underlining.

sonii strain used in this study which is also considered to be a severe strain. Three of these changes are substitutions whereas the fourth is a deletion. For that reason, this PSTV strain is 1 nucleotide shorter and encompasses only 358 nucleotides. It is interesting to note that the AA → U exchange at position 120 is also found in the previously determined sequence of a mild strain (12). This mild strain contains two other differences which are located around position 309. One is an exchange (A → U at position 309), the other is an insertion of U at position 310-311. In the *S. commersonii* strain, at positions 309-310 an AC → CU exchange has been noted (see Fig. 5). Thus, all three PSTV strains differ only within two regions which are located at both sites of the highly conserved central loop of the secondary structure model (2-6). In one region, the *S. commersonii* strain shares some nucleotide differences with the mild strain whereas in the other region exchanges have been found in all three strains. At least in that part of the sequence, the numbers of nucleotides of the two severe strains agree with each other.

The question whether these nucleotides contribute in some way to the pathogenic response in the host remains to be elucidated. More strains have to be isolated to correlate nucleotide differences with the secondary structure of the RNA and its pathogenicity. If the complete DNA copy of PSTV described

here would produce biologically functional viroid RNA molecules in plant cells, as has been found for cloned copies of phage Q β and poliovirus RNAs in their hosts (13, 14), it might be a powerful tool in generating PSTV mutants. Furthermore, plasmid pAV401 can perfectly well be used as a hybridization probe for the detection of potato spindle tuber viroid in potato tubers (10).

ACKNOWLEDGEMENTS

We thank Willem Mosch for the gift of PSTV RNA, Piet Madern for art work and photography, Pim Zabel and Rob Goldbach for critical reading the manuscript and Marie-José van Neerven for typing the manuscript. This work was supported by the Research Institute for Plant Protection (I.P.O.). Peter van Wezenbeek is a postdoctoral fellow of the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

*Present address: Laboratory of Biochemistry, University of Nijmegen, 6525 EZ Nijmegen, The Netherlands.

[†]Gorlaeus Laboratories, State University of Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands.

§To whom correspondence should be addressed.

REFERENCES

1. Diener, T.O. (1979). *Science* 205, 859-866.
2. Gross, H.J., Domdey, H., Lossow, C., Jank, P., Raba, M., Alberty, H. and Sanger, H.L. (1978). *Nature* 273, 203-208.
3. Haseloff, J. and Symons, R.H. (1981). *Nucl. Acids Res.* 9, 2741-2752.
4. Visvader, J.E., Gould, A.R., Bruening, G.E. and Symons, R.H. (1982). *FEBS Lett.* 137, 288-292.
5. Symons, R.H. (1981). *Nucl. Acids Res.* 9, 6527-6537.
6. Haseloff, J., Mohamed, N.A. and Symons, R.H. (1982). *Nature* 299, 316-321.
7. Diener, T.O. (1979). *Viroids and Viroid Diseases*. (Wiley, New York).
8. Morris, T.J. and Smith, E.M. (1977). *Phytopathology* 67, 145-150.
9. Mosch, W.H.M., Huttinga, H., Hakkaart, F.A. and de Bokx, J.A. (1978). *Neth.J.Pl.Path.* 84, 85-93.
10. Owens, R.A. and Diener, T.O. (1981). *Science* 213, 670-672.
11. Dickson, E., Robertson, H.D., Niblett, C.L., Horst, R.K. and Zaitlin, M. (1979). *Nature* 277, 60-62.
12. Gross, H.J., Liebl, U., Alberty, H., Krupp, G., Domdey, H., Ramm, K. and Sanger, H.L. (1981). *Biosci.Rep.* 1, 235-241.
13. Taniguchi, T., Palmieri, M. and Weissmann, C. (1978). *Nature* 274, 223-228.
14. Racaniello, V.R. and Baltimore, D. (1981). *Science* 214, 916-919.
15. Wilson, G.G. and Murray, N.E. (1979). *J.Mol.Biol.* 132, 471-491.
16. Panet, A., van de Sande, J.H., Loewen, P.C., Khorana, H.G., Raaij, A.J., Lillehaug, J.R. and Kleppe, K. (1973). *Biochemistry* 12, 5045-5050.
17. Gammack, R.H. and Harris, P.S. (1973). *EPPO-Bull.* 3, 117-118.
18. Diener, T.O., Hadidi, A. and Owens, R.A. (1977). In: *Methods in Virology*, Maramorosch, K. and Koprowski, H., Eds. (Academic Press, New York)

- 6, 185-217.
19. Sanger, H.L., Ramm, K., Domdey, H., Gross, H.J., Henco, K. and Riesner, D. (1979). FEBS Lett. 99, 117-122.
20. Godson, G.N. and Vapnek, D. (1973). Biochim.Biophys.Acta 299, 516-520.
21. Van der Marel, G.A., Wille, G., Marugg, J.E., De Vroom, E., Tromp, M., Van Boeckel, C.A.A. and Van Boom, J.H. (1982). Rec.Trav.Chim.Pays-Bas, in press.
22. Messing, J., Crea, R. and Seeburg, P.H. (1981). Nucl.Acids Res. 9, 309-321.
23. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc.Natl.Acad.Sci.USA 74, 5463-5467.
24. Messing, J. (1982). In: Genetic Engineering, Principles and Methods, Hollaender, A. and Setlow, J., Eds. (Plenum, New York) 4, in press.
25. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980). J.Mol.Biol. 143, 161-178.
26. Rohde, W., Schnolzer, M., Rackwitz, H.-R., Haas, B., Seliger, H. and Sanger, H.L. (1981). Eur.J.Biochem. 118, 151-157.
27. Van Wezenbeek, P.M.G.F., Hulsebos, T.J.M. and Schoenmakers, J.G.G. (1980). Gene 11, 129-148.
28. Volckaert, G., Tavernier, J., Derynck, R., Devos, R. and Fiers, W. (1981). Gene 15, 215-223.
29. Sutcliffe, J.G. (1979). Cold Spring Harbor Symp. Quant. Biol. 43, 77-90.