Volume 17 Number 6 1989 Nucleic Acids Research

Pea early browning virus RNA1 encodes four polypeptides including a putative zinc-finger protein

S.A.MacFarlane, S.C.Taylor, D.I.King, G.Hughes* and J.W.Davies

John Innes Institute and AFRC Institute of Plant Science Research, Colney Lane, Norwich NR4 7UH, UK

Received January 20, 1989; Revised and Accepted February 21, 1989 EMBL accession no. X14006

ABSTRACT

We have determined the complete nucleotide sequence of RNA1 of the tobravirus pea early browning virus [PEBV] from an overlapping series of cDNA clones. The 7073 nucleotide sequence contains four open reading frames [ORFs]. The 5′ proximal ORF encodes a 141K polypeptide, and readthrough of the opal [UGA] termination codon of this ORF would lead to the synthesis of a second, 201K polypeptide. Both of these polypeptides have extensive amino acid homology with the putative replicase proteins of tobacco rattle virus [TRV] and tobacco mosaic virus [TMV]. The third ORF encodes a 30K polypeptide which has homology with the TRV 29K and TMV 30K putative cell-to-cell spread proteins. The fourth, 3′ proximal ORF encodes a 12K polypeptide which has extensive homology with the TRV 16K protein whose function is unknown. Examination of the amino acid sequences of the 12K and 16K gene products reveals in each the presence of two multiple-cysteine/histidine motifs, a finding which suggests that these proteins might have zinc and/or nucleic acid-binding properties.

INTRODUCTION

Pea early browning virus [PEBV] is a member of the tobravirus group of plant viruses. These viruses have a genome consisting of two positive-strand RNA molecules, each encapsidated separately in a rod-shaped particle (1). The tobraviruses are divided into three subgroups based on serological properties: these are PEBV, tobacco rattle virus [TRV], and pepper ringspot virus [PRV, formerly the CAM strain of TRV] (2,3). These viruses are transmitted in the field by soil-inhabiting nematodes and are known to infect a very wide range of plant species.

The RNA1 molecules of all three viruses are of similar length whilst the size of RNA2 differs markedly from isolate to isolate (4,5,6). Polyacrylamide gel electrophoresis studies estimated

that the British strain [SP5] of PEBV had an RNA1 of approximately 7300 nucleotides and an RNA2 of about 3800 nucleotides (7). Since RNA1 can systemically infect plants in the absence of RNA2, the larger RNA must encode both replication and virus-movement functions. In vitro translation of PEBV SP5 RNA1 produced two large polypeptides with apparent molecular weights of 165K and 134K, and it was suggested that production of the larger polypeptide resulted from readthrough of the 134K polypeptide termination codon (8). Analysis of cDNA sequence encoding the analogous proteins in TRV strain SYM has identified regions sharing amino acid homology with viral RNA-dependent RNA polymerases (9,10). TRV-SYM RNA1 has also been shown to encode two other proteins, both expressed from subgenomic RNAs (11), one of which has homology to the TMV 30K cell-to-cell spread protein. Neither of these proteins nor their subgenomic RNAs have, as yet, been found associated with PEBV.

The biology of PEBV differs from TRV in its adaptation to leguminous plant hosts and its spectrum of nematode vectors (12,13). It is anticipated that the elucidation of the nucleotide sequence of PEBV RNA1 and its comparison with the sequences obtained from a number of strains of TRV will be an important step in understanding this particular group of plant viruses.

METHODS

Isolation of virus. The SP5 isolate of the British strain of PEBV [inoculum provided by D.Robinson and M.Mayo, SCRI, Invergowrie] was propagated in Nicotiana clevelandii. Infected plant tissue was homogenised in 30mM sodium phosphate/0.15% [v/v] thioglycollic acid, pH7.5 and frozen overnight at -20°C. The homogenate was thawed to room temperature and clarified by filtration through muslin followed by centrifugation at 6000 r.p.m. for 5 min. at 4°C. Virus particles were precipitated by the addition of polyethylene glycol 6000 [to a final concentration of 10% w/v in 2% NaCl], collected by centrifugation [10K r.p.m.,5 min.,4°C] and resuspended in 30mM phosphate, pH7.5. Any remaining plant material was removed by repeated extractions with a 1:1 mixture of chloroform/iso-butanol and further rounds of low speed centrifugation. The supernatant was centrifuged at

high speed [40K r.p.m.,80 min.,4°C], the pellet resuspended in 10mM Tris-Cl, pH7.5 and then centrifuged through a 30% sucrose cushion [29K r.p.m.,150 min.,15°C]. The pellet was finally resuspended in 10mM Tris-Cl, pH7.5.

Preparation of RNA. Purified virus was incubated, for 2 hr. on ice, with 25µg/ml RNAse-free DNAsel [Sigma], followed by the addition of Proteinase K [Sigma, 300µg/ml] and a further incubation at 37°C for 2 hr. RNA was extracted as described previously (8), and RNAs 1 and 2 were separated by centrifugation through a 15-30% [w/v] sucrose gradient [30K r.p.m., 12 hr., 20°C in an SW41 rotor].

cDNA synthesis and cloning. The determination of the 3′ terminal base and the subsequent sequencing of a portion of the 3′ non-coding region of PEBV RNA1 was described previously (14). Having obtained this information, a 10-base oligonucleotide complementary to the 3′end of RNA1 was used to initiate first strand cDNA synthesis. Later a 19-base oligonucleotide [sequence derived from 3′cDNA clones] was used. cDNA synthesis was carried out as described before (15), using AMV reverse transcriptase (Pharmacia). cDNA clones were also obtained by making use of a commercial cDNA synthesis kit [Amersham]. The double-stranded cDNA was blunt-ended with T4 DNA polymerase and/or digested with restriction enzymes and cloned into Bluescript [Stratagene] for subsequent sequencing.

DNA sequencing. DNA sequencing was carried out using the dideoxy chain-termination method (16) with α^{35}S-dATP as the labelled nucleotide (17). Regions of ambiguity were resolved by using the Sequenase DNA sequencing kit [USB corporation]. One cDNA clone, number 191, was found to be lacking nucleotide 399 which was included in the final sequence [figure 1]. The correct sequence was confirmed by direct RNA sequencing of this region using a method described previously (18).

Determination of 5′ terminal sequence. cDNA corresponding to the 5′ end of PEBV RNA1 was synthesised using a kinase-labelled oligonucleotide primer complementary to nucleotides 121-137 of the finally derived sequence. The primer-extension product was purified on an 8% polyacrylamide gel and sequenced by the chemical method (19). The sequence was confirmed by repeating the

procedure with a second primer, hybridising closer to the 5´end [nucleotides 44-62].

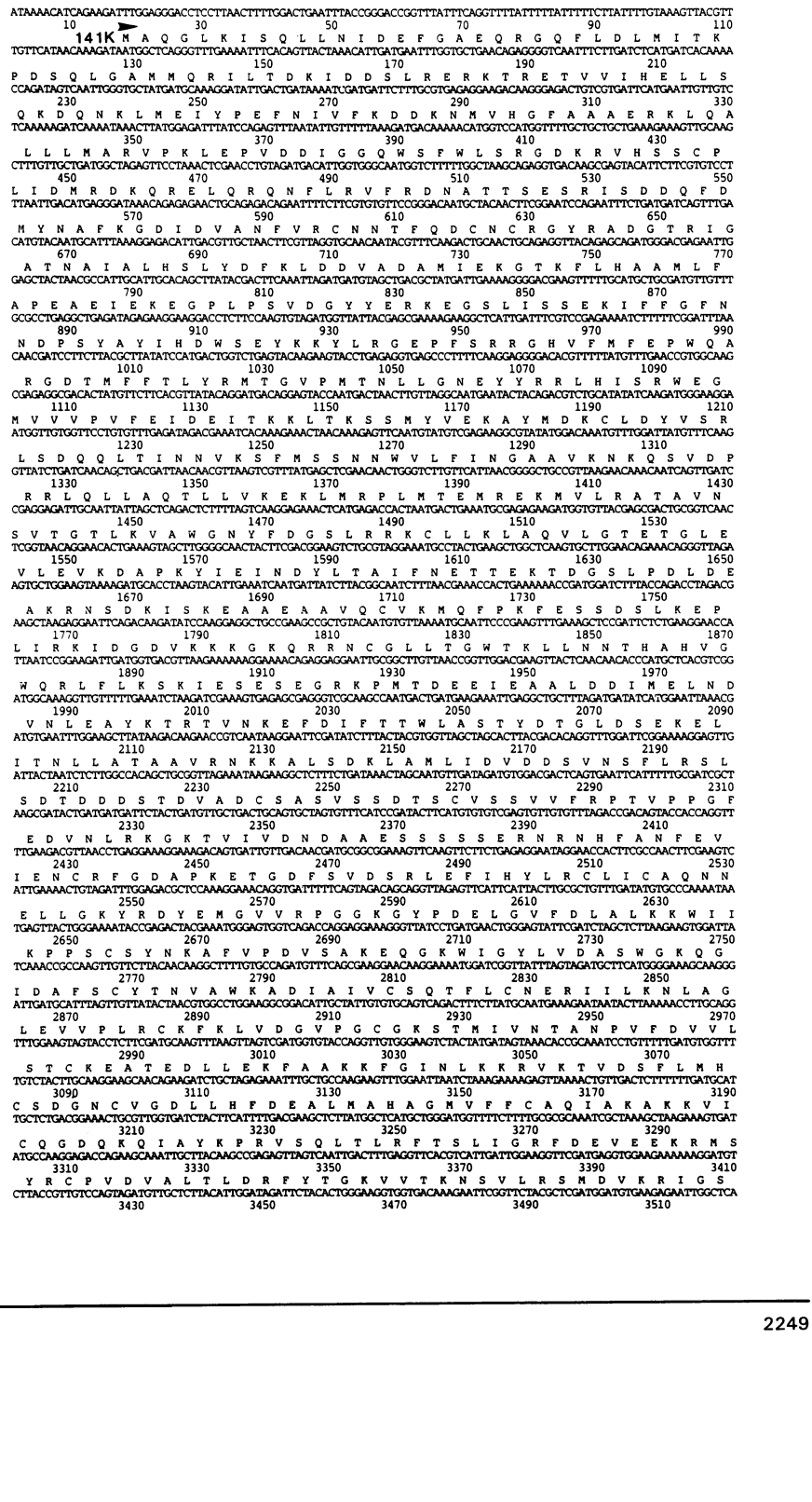
Computing. Assembly and analysis of the cDNA sequence was carried out using programmes described previously (20,21,22,23)

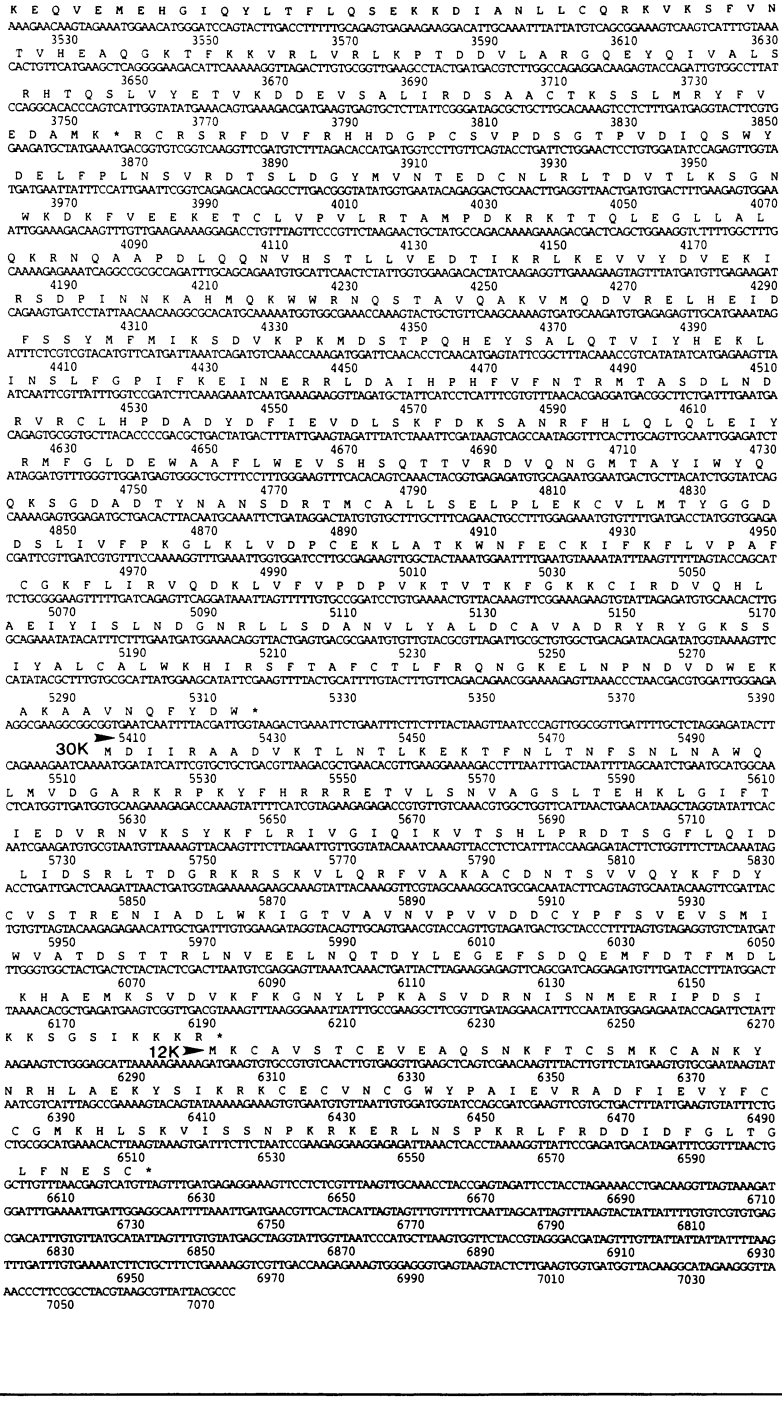
RESULTS AND DISCUSSION

Sequence was derived from an overlapping series of 25 cDNA clones extending from nucleotide 83 to the extreme 3´ end of the RNA. Seventy percent of the cDNA was represented by two or more independently obtained clones. The complete cDNA was sequenced on both strands and, on average, each base was sequenced 5.3 times. The sequence from the 5´ terminus to position 83 was obtained by primer extension as described above. The sequence of a complete DNA version of PEBV RNA1 is presented in figure 1.

PEBV RNA1 is 7073 nucleotides in length. The overall base composition is: U 29.0%, C 15.8%, A 30.6%, G 24.6%. Base heterogeneities were found at five positions in the cDNA sequence. Both G and A were found at position 191 giving the triplets GGT [encoding glycine] and AGT [encoding serine], C and T were at position 406 giving GTC and GTT [both encoding glycine], C and T were found at position 1190 giving CAT [encoding histidine] or TAT [encoding tyrosine], nucleotide 4420 was either T or C giving the triplets TTC and TTT [both encoding phenylalanine] and either T or C was at position 6626 in the 3´ non-coding region.

Computer analysis revealed four open reading frames [ORFs] on the virion-sense RNA [figure 2]. The first ORF starts at position 128 [the first AUG codon] and extends to position 3865. Its 1246 amino acid [Aa] translation product has a calculated molecular weight [mol.wt.] of 141,383 [141K] daltons. Readthrough of the 141K protein opal [UGA] termination codon extends translation to the end of the second ORF [position 5425] resulting in a 1765 residue protein with a mol.wt. of 201,600 [201K]. ORF3 extends from positions 5515 to 6300 and has a 262 Aa translation product with a mol.wt. of 30,119 [30K]. ORF4 extends from positions 6300 to 6620 and has a 107 Aa translation product with a mol.wt. of 12,363 [12K].




30K
12K

Non-coding regions

5´ Leader sequence. Upstream of the first ORF is a non-coding region of 127 bases containing no AUG codons. In contrast to the direct repeats identified in the leader sequences of a number of strains of TRV (10), no such sequences were found in the 5´ non-coding region of PEBV RNA1. The leader region is, however, very A/U rich :88/127 nucleotides [69%] as compared to a figure of 59% for the entire genomic RNA.

The first nine nucleotides at the extreme 5´ terminus of PEBV RNA1 [ATAAAACAT] are identical to those found at the 5´ termini of the RNA1 molecules of TRV strains SYM, ORY, PRN, N5 and the RNA2 molecules of TRV strains SYM, ORY, PRN, PSG and TCM (10). Experiments combining RNA1 and RNA2 molecules from different sources have shown that pseudo-recombinants can be successfully generated by mixing RNA1 and RNA2 from different strains of TRV, whereas combinations of RNAs from TRV and PEBV are not viable (24). The identity of the 5´ termini of both of these viruses would seem to exclude the direct involvement of these nucleotides in the selection of template by the virus-encoded RNA polymerase.

A comparison of the sequences around the AUG translation initiation codons of the 141K, 30K and 12K genes is presented in table 1. The sequence flanking the 141K gene AUG most closely resembles the consensus sequence identified for plant genes (25), and it is probable that this gene is translated with greater efficiency than are the 30K and 12K genes. A second in-frame AUG codon is present 60 nucleotides downstream from the AUG proposed to initiate translation of the 12K gene. However, the second AUG is in no better a context than the first and is not present in the TRV 16K gene making it likely that translation of the 12K gene does indeed initiate at the first AUG at position 6300.

Intergenic regions. The region between the 201K and 30K genes comprises 86 bases. This compares with only a single base for the same region in TRV-SYM and TRV-PSG (9,26). In TMV these genes overlap by 17 bases (27). Conversely, the region between

Figure 1. Nucleotide sequence of a complete DNA copy of PEBV RNA1. The ATG initiation codon of each ORF is highlighted with an arrow. The approximate molecular weights of the translation products of the ORFs are marked.

Table 1. Alignment of sequences flanking the translation initiation codons of the PEBV 141K, 30K and 12K genes [a].

Gene	position	sequence
141K	124	GATA**ATG**GC
30K	5511	CAAA**ATG**GA
12K	6296	AAAG**ATG**AA
plant consensus (ref.25)		AACA**ATG**GC

[a]Initiation codon highlighted in bold, other sequences in common with consensus underlined.

the 29K gene and 16K gene is 24 bases in TRV, whereas the UGA termination codon of the PEBV 30K gene overlaps the AUG initiation codon of the 12K gene.

3′ Non-coding region. PEBV RNA1 has a 3′ non-coding region of 450 nucleotides. This compares with 255 nucleotides for TRV-SYM, -PSG and -TCM RNA1 (9,26,28). A twelve nucleotide sequence at the extreme 3′ terminus [GTTATTACGCCC] is also present in all strains of TRV and PRV analysed to date. In addition, a comparison using the computer programme Diagon shows that there is considerable homology between the 3′ non-coding regions of PEBV and TRV [figure 3]. This finding differs with the results of previous hybridisation studies which have suggested that, even in this region, there is little or no nucleotide sequence homology between TRV and PEBV (3,29).

It is known that aspects of both the nucleotide sequence and secondary structure at the 3′termini of plant virus RNAs affect their replication (30,31,32). RNA2 from TRV strain PSG has been shown to adopt an extensively folded secondary structure with two

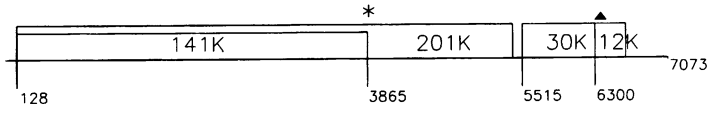


Figure 2. Genome organisation of PEBV RNA1. Numbers below the figure are nucleotides, boxed numbers are the approximate molecular weights of the polypeptides encoded by RNA1 ORFs. The position of the readthrough termination codon is shown by an asterisk. The triangle marks the location within the 12K polypeptide of the putative zinc-finger domains.

large hairpin loops [between 130-45 bases from the 3´ end] and a pseudoknot [encompassing 40 bases] at the extreme 3´ end (33). The pseudoknotted region resembles the aminoacyl acceptor arm of tRNA and can be adenylated *in vitro* by nucleotidyl transferase but cannot be aminoacylated [in contrast to a number of other plant virus RNAs] (34). PEBV RNA1 also has an identifiable pseudoknot at its 3´ end, although its helical, double-stranded RNA domain is one basepair longer than that of TRV.

Coding regions

Dot matrix comparisons emphasise the high degree of homology apparent between the proteins of PEBV, TRV-SYM and TMV [figures 4a-c].

141/201K putative replicase proteins. The PEBV 141K and 201K proteins very closely resemble the 134K and 194K replicase proteins of TRV, and to a lesser extent the 126K and 183K proteins of TMV [figures 4a, 4b]. There is extensive amino acid homology between the NH_2-proximal regions of the PEBV 141K and TRV 134K proteins, between the COOH-proximal regions of the 141K and 134K proteins, and also between the regions extending from the 141K/134K termination codons to the COOH-termini of the PEBV 201K and TRV 194K proteins. Consequently, there are three regions of extensive homology between the PEBV and TMV replicase proteins which correspond to those that have been previously described for TRV and TMV (10).

The COOH-terminal region of the 141K protein [encoded by nucleotides 3010 to 3750] contains the six conserved motifs identified as belonging to a large superfamily of nucleic acid replication and/or recombination proteins (35,36,37). The first of these motifs [VDGVPGCGKSTMIV at position 3010 in PEBV] comprises the nucleotide-binding fold found in ATP and GTP-binding proteins (38,39). The functions of the five remaining conserved motifs are not known. The COOH-terminal region of the 201K protein contains a second set of four conserved motifs [between nucleotides 4660 to 5080] which includes the -GDD- "box" common to RNA-dependent RNA polymerases (40,41).

It is interesting to note that the only region disimilar to TRV lies between residues 475-719 [nucleotides 1553 to 2285] in

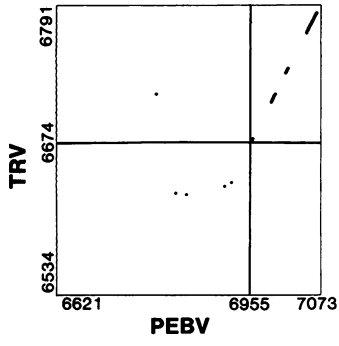
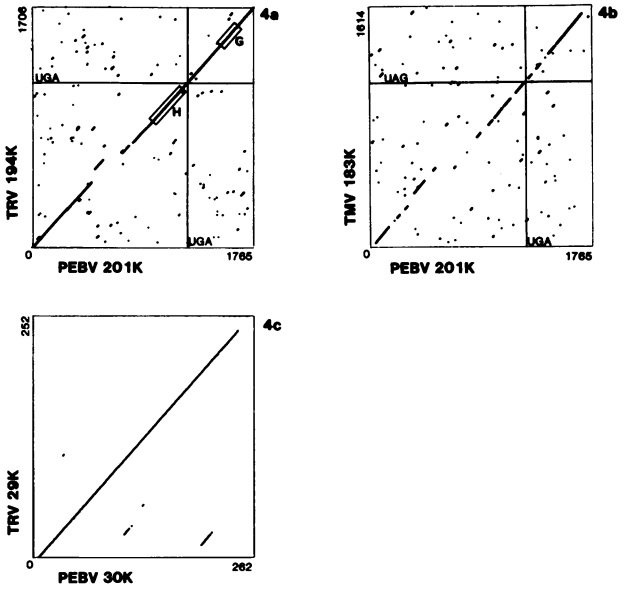


Figure 3. Diagon plot of 3′ non-coding regions of PEBV and TRV-SYM RNA1. Numbers indicate nucleotides (relative to 5′ end), cross-hairs indicate start of homology. For comparison span length was 11 and number of identical, matched bases was 8.

the PEBV 141K protein, and it is possible that the differences in this region are an important factor in the inability *in vivo* of PEBV-encoded polymerase to support the replication of TRV genomic RNA and *vice versa* (24).

30K putative cell-to-cell spread protein. The PEBV 30K protein is slightly larger than its TRV counterpart [262 Aa compared to 252 Aa]. The Diagon plot [figure 4c] shows that these two proteins are highly homologous throughout their entire lengths, excepting their extreme COOH-termini.The TMV 30K protein has been implicated in cell-to-cell spread (42,43,44). Amino acid homology between the TMV 30K protein and the TRV 29K protein has been noted before (9). Our data extends this homology to the PEBV 30K protein.

12K protein: a putative zinc-finger protein . A computer search of the NBRF protein sequence databank showed that the PEBV 12K protein had significant amino acid homology only with the TRV 16K protein. Alignment of these two proteins using the UWGCG GAP programme [figure 5] reveals that the difference in the sizes of these two proteins results from a continuous span of 35 amino acids absent from the PEBV protein but present in the TRV protein. Closer examination of the amino acid sequences of these proteins reveals the presence in each of a duplicated motif of cysteine and histidine residues. The motifs of the two proteins differ from one another by the substitution of histidine for



Figures 4a-c. Diagon plots of the predicted amino acid sequences of a) PEBV 201K against TRV-SYM 194K, b) PEBV 201K against TMV 183K, c) PEBV 30K against TRV-SYM 29K. Positions of UAG and UGA termination codons are indicated by cross-hairs, location in PEBV 201K of conserved "helicase" and "GDD-box" motifs are shown by H and G respectively. For amino acid comparisons span length was 31 and set proportional score was 340.

cysteine in the first motif of TRV, and also by the addition of an extra residue between the cysteine/histidine pairs of the first motif in TRV.

Such an arrangement of amino acids is consistent with the pattern of conserved cysteine and histidine residues found in "zinc-finger" proteins (45,46). These proteins, typified by the well-studied Xenopus transcription factor IIIA, are characterised by their possession of one or more domains containing two cysteine and two histidine residues arranged in a regular pattern, with a conserved aromatic residue [phenylalanine or tyrosine] and a leucine residue between the Cys/Cys and His/His pairs. (47). It is now becoming clear that proteins containing multiple-cysteine motifs yet lacking the rigidly conserved sequence of amino acids identified in the TFIIIA and

```
PEBV       [1]  MKCAVSTCEVE.AQSNKFTCSMKCANKYNRHLEAKYSIKRKCECVNCG
                : : :  :  :       :::: :::  ::::: :: :::    ::
TRV (SYM)  [1]  MTCVLKGCVNEVTVLGHETCSIGHANKLRKQVADMVGVTRRCAENNCG

TRV (PSG)  [1]  MTCVLKGCVNEVTVLGHETCSIGHANKLRKQVADMVGVTRRCAENNCG

TRV (TCM)  [1]  MTCVLKGCVNEVTVLGHETCSIGHANKLRKQVADMVGVTRRCAENNCG


PEBV            WYPAIEVRADFIEVYFCCGMKHL.......................
                ::  :      :::: :::  ::
TRV (SYM)       WFVCVVINDFTFDVYNCCGRSHLEKCRKRVETRNREIWKQIRRNQAEN
                   **                          *           *   *
TRV (PSG)       WFVCIIINDFTFDVYNCCGRSHLEKCRKRVEARNREIWKQIRRIQAES
                   *                           *          **  ** *
TRV (TCM)       WFVCIIINDFTFDVYNCCGRSHLEKCRKRFEARNREIWKQVERIRGEK


PEBV            ..........SKVISSNPKRKERLNSPKRLFRDDIDFGLTGLFNESC
                          ::      :    :  ::::::::: :::  ::
TRV (SYM)       MSATAKKSHNSKTSKKKFKEDREFGTPKRFLRDDVPFGIDRLFAF
                *   *       *           *          *   *  *
TRV (PSG)       SSATRKKSHNSKNSKKKFKEDREFGAPKRFLRDDVPLGIDQLFVF
                *   *   *  *      ***  *
TRV (TCM)       ASATVKKSHKSPKSKKKFKERKDFGTPKRFLRDDVPLGIDQLFVF
```

Figure 5. Alignment of tobraviral 12K and 16K putative zinc-finger proteins. PEBV protein encoded by nucleotides 6300-6620 (figure 1). TRV sequences see refs. 9,26,28. Conserved cysteine/histidine residues of putative finger motifs are highlighted and underlined. Numbers in brackets are amino acid residues, * indicates differences between adjacent residues,: indicates conserved residues between PEBV and TRV, . indicates gaps inserted to optimise homologies.

related finger proteins are also capable of zinc-dependent DNA-binding (48,49).

A sequence encoding a Cys X_2 Cys X_4 His X_4 Cys motif has been identified in the virion-associated nucleic acid-binding proteins [NBPs] of mammalian retroviruses. Site-directed mutagenesis of the Rous sarcoma virus NBP affects viral RNA packaging, dimerisation and the correct positioning of a tRNA primer onto the genomic RNA prior to virus replication (50,51).
Such a motif is also present in the coat protein gene [ORF IV] of cauliflower mosaic virus [CaMV] (52). Although this plant virus encapsidates double-stranded DNA, it nevertheless replicates by reverse transcription of an RNA intermediate.

A more convincing argument for the existence of zinc-binding proteins in plant viruses comes from recent studies of the ilarvirus, tobacco streak virus [TSV]. Replication of the virus requires that a few copies of the coat protein bind to the 3´end of the genomic RNA. A three-cysteine one-histidine motif is present in the amino terminal region of the coat protein and,

moreover, substantial amounts of zinc are contained within the virus capsid (53).

The PEBV 12K protein can be divided into two domains : an N-terminal domain containing two cysteine-histidine motifs and a C-terminal domain containing two proline residues and eight basic [lysine or arginine] residues. Homologous domains are present in the TRV 16K protein. This arrangement is similar to that of the retroviral finger-proteins which have three identifiable domains: an N-terminal basic domain, a central "Cys-His box", and a C-terminal domain rich in proline and basic residues (54).

A functional importance for the putative tobraviral zinc fingers is suggested by a comparison of the amino acid sequences of the 16K proteins from three different strains of TRV [figure 5]. Of the 27 differences occuring between these proteins, 24 are in the COOH-terminal domain while only 3 are in the zinc finger domain [none of which include the cysteine and histidine residues]. As was discussed above PEBV and TRV possess tRNA-like structures at their 3´ends (33), which, in a manner analogous to the situation found with RSV and MuLV, might interact with the p12 and p16 proteins respectively.

The striking similarity in genome organisation of PEBV and TRV can be extended to include TMV, with the exceptions of an absence of a 12K/16K homologue in TMV and the displacement of the PEBV and TRV coat protein genes to a separate RNA [RNA2]. Coat protein-deficient mutants of TMV produce lesions by cell-to-cell spread within the infected leaf but do not spread into the upper, uninoculated leaves (55). In contrast, NM-infections of PEBV, comprising naked RNA1 only, are quite commonly found. Although the spread of RNA1 in these plants is considered to occur by passage between adjacent cells, the infection does become systemic and eventually reaches the upper leaves. Our identification of two putative zinc-fingers in the 12K and 16K proteins leads us to suggest that they may bind to one or both of the viral RNAs. This interaction might protect the RNA from degradation by host ribonucleases and thus extend the range of spread of the RNA in the infected plant. Alternatively, analogies with the retroviral nucleic acid-binding proteins might suggest that the 12K protein could be involved in encapsidation of the

genomic RNAs by the coat protein. If mutations in the 12K gene were to affect virus particle assembly, subsequent transmission by nematodes might also be disrupted.

ACKNOWLEDGEMENTS

The authors would like to thank Dr.G.Lomonossoff for his considerable help and advice during this work. We also thank M.Shanks for assistance with both the cDNA cloning and sequencing methodologies. M.G.Goulden is thanked for isolating and sequencing a 1.0 kb cDNA clone corresponding to part of the 201K and 30K genes, and communicating to us the method for direct RNA sequencing. Dr.S.N.Covey is acknowledged for helpful discussions, and Dr.R.Hull for critical reading of this manuscript.

*Present address: Department of Veterinary Microbiology, Western College of Veterinary Medicine, University of Saskatchewan, Saskatchewan S7N 0WO, Canada

REFERENCES

1. Harrison,B.D. and Robinson,D.J. (1986) Tobraviruses. In The Plant Viruses, vol. 2, pp.339-369. Edited by M.H.V.van Regenmortel and H.Fraenkel-Conrat, New York: Plenum Press.
2. Harrison,B.D. and Robinson,D.J. (1978) Adv.Virus Res., **23**, 25-77.
3. Robinson,D.J. and Harrison,B.D. (1985) J.Gen.Virol., **66**, 171-176.
4. Robinson,D.J. (1983) J.Gen.Virol., **64**, 657-665.
5. Pelham,H.R.B. (1979) Virology, **97**, 256-265.
6. Harrison,B.D. and Woods,R.D. (1966) Virology, **28**, 610-620.
7. Cooper,J.I. and Mayo,M.A. (1972) J.Gen.Virol., **16**, 285-297.
8. Hughes,G., Davies,J.W. and Wood,K.R. (1986) J.Gen.Virol., **67**, 2125-2133.
9. Boccara,M., Hamilton,W.D.O. and Baulcombe,D.C. (1986) EMBO J., **5**, 223- 229.
10. Hamilton,W.D.O., Boccara,M., Robinson,D.J. and Baulcombe,D.C. (1987) J.Gen.Virol., **68**, 2563-2575.
11. Robinson,D.J., Mayo,M.A., Fritsch,C., Jones,A.T. and Raschke,J.H. (1983) J.Gen.Virol., **64**, 1591-1599.
12. van Hoof,H.A., Maat,D.Z. and Seinhorst,J.W. (1966) Neth.J.Pl.Path., **72**, 253-258.
13. Harrison,B.D. (1966) Ann.Appl.Biol., **57**, 121-129.
14. Hughes,G. (1985) Ph.D. thesis, University of Birmingham.
15. Lomonossoff,G.P., Shanks,M., Matthes,H.D., Singh,M. and Gait,M.J. (1982) Nucleic Acids Res., **10**, 4861-4872.
16. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc.Natl.Acad.Sci.USA, **74**, 5463-5467.
17. Biggin,M.D., Gibson,T.J. and Hong,G.F. (1983)

Proc.Natl.Acad.Sci. USA, **80**, 3963-3965.
18. Meshi,T., Ishikawa,M., Takamatsu,N., Ohno,T. and Okada,Y. (1983) FEBS Lett., **162**, 282-285.
19. Maxam,A.M. and Gilbert,W. (1977) Proc.Natl.Acad.Sci.USA, **74**, 560-564.
20. Devereux,J., Haeberli,P. and Smithies,O. (1984) Nucleic Acids Res., **12**, 387-395.
21. Staden,R. (1980) Nucleic Acids Res., **8**, 3673-3694.
22. Staden,R. (1984) Nucleic Acids Res., **12**, 521-538.
23. Lipman,D.J. and Pearson,W.R. (1985) Science, **227**, 1435-1551.
24. Robinson,D. and Harrison,B. (1984) Abstracts, 6th International Congress of Virology, Sendai, p.207.
25. Lutcke,H.A., Chow,K.C., Mickel,F.S., Moss,K.A., Kern,H.F. and Scheele,G.A. (1987) EMBO J., **6**, 43-48.
26. Cornelissen,B.J.C., Linthorst,H.J.B., Brederode,F.T. and Bol,J.F. (1986) Nucleic Acids Res., **14**, 2157-2169.
27. Goelet,P., Lomonossoff,G.P., Butler,P.J.G., Akam,M.E., Gait,M.J. and Karn,J. (1982) Proc.Natl.Acad.Sci.USA, **79**, 5818-5822.
28. Angenent,G.C., Linthorst,H.J.B., van Belkum,A.A., Cornelissen,B.J.C. and Bol,J.F. (1986) Nucleic Acids Res., **14**, 4673-4682.
29. Robinson,D.J., Hamilton,W.D.O., Harrison,B.D. and Baulcombe,D.C. (1987) J.Gen Virol., **68**, 2551-2561.
30. Bujarski,J.J., Ahlquist,P., Hall,T.C., Dreher,T.W. and Kaesberg,P. (1986) EMBO J., **5**, 1769-1774.
31. Bujarski,J.J., Dreher,T.W. and Hall,T.C. (1985) Proc.Natl.Acad.Sci.USA. **82**, 5636-5640.
32. Miller,W.A., Bujarski,J.J., Dreher,T.W. and Hall,T.C. (1986) J.Mol.Biol., **187**, 537-546.
33. van Belkum,A., Cornelissen,B., Linthorst,H., Bol,J., Pley,C. and Bosch,L. (1987) Nucleic Acids Res., **15**,2837-2850.
34. Haenni, A.-L., Joshi,S. and Chapeville,F. (1982) Prog.Nucleic Acid Res.Mol.Biol., **27**, 85-104.
35. Hodgman,T.C. (1988) Nature, **333**, 22-23.
36. Gorbalenya,A.E., Koonin,E.V., Donchenko,A.P. and Blinov,V.M. (1988) Nature, **333**, 22.
37. Lane.D. (1988) Nature, **334**, 478.
38. Walker,J.E., Saraste,M., Runswick,M.J. and Gay,N.J. (1982) EMBO J., **8**, 945-951.
39. Evans,R.K., Haley,B.E. and Roth,D.A. (1985) J.Biol.Chem., **260**, 7800-7804.
40. Kamer,G. and Argos,P. (1984) Nucleic Acids Res., **12**, 7269-7282.
41. Gorbalenya,A.E. and Koonin,E.V. (1988) Nucleic Acids Res., **16**, 7734.
42. Leonard,D.A. and Zaitlin,M. (1982) Virology, **117**, 416-424.
43. Deom,C.M., Oliver,M.J. and Beachy,R.N. (1987) Science, **237**, 389-394.
44. Meshi,T., Watanabe,Y., Saito,T., Sugimoto,A., Maeda,T. and Okada,Y. (1987) EMBO J., **6**, 2557-2563.
45. Klug,A. and Rhodes,D. (1987) Trends Biochem.Sci., **12**, 464-469.

46. Berg,J.M. (1986) Science, **232**, 485-487.
47. Evans,R.M. and Hollenberg,S.M. (1988) Cell, **52**, 1-3.
48. Freedman,L.P., Luisi,B.F., Korszun,Z.R., Basavappa,R., Sigler,P.B. and Yamamoto,K.R. (1988) Nature, **334**, 543-546.
49. Severne,Y., Wieland,S., Schaffner,W. and Rusconi,S. (1988) EMBO J., **7**, 2503-2508.
50. Prats,A.C., Sarih,L., Gabus,C., Litvak,S., Keith,G. and Darlix,J.L. (1988) EMBO J., **7**, 1777-1783.
51. Meric,C., Gouilloud,E. and Spahr,P.-F. (1988) J.Virology, **62**, 3328-3333.
52. Covey,S.N. (1986) Nucleic Acids Res., **14**, 623-633.
53. Sehnke,P.C., Mason,A.M., Hood,S.J., Lister,R.M. and Johnson,J.E. (1989) Virology, **168**, 48-56..
54. Schwartz,D.E., Tizard,R. and Gilbert,W. (1983) Cell, **32**, 853-869.
55. Takamatsu,N., Ishikawa,M., Meshi,T. and Okada,Y. (1987) EMBO J., **6**, 307-311.