Nucleotide sequences of coat protein genes for three isolates of barley yellow dwarf virus and their relationships to other luteovirus coat protein sequences

J. R. Vincent,* P. P. Ueng,† R. M. Lister and B. A. Larkins‡

Department of Botany and Plant Pathology, Lilly Hall of Life Sciences, Purdue University, West Lafayette, Indiana 47907, U.S.A.

Barley yellow dwarf virus (BYDV) can be separated into two groups based on, among other criteria, serological relationships that are presumably governed by the viral capsid structure. Nucleotide sequences for the coding regions of coat proteins of approximately 22 K were identified for the MAV-PS1, P-PAV (group 1) and NY-RPV (group 2) isolates of BYDV. The MAV-PS1 and P-PAV coat protein sequences shared 71% deduced amino acid similarity whereas that of the NY-RPV isolate shared no more than 51% similarity with either the MAV-PS1 or the P-PAV sequence. Other comparisons showed that these and other BYDV coat protein sequences examined to date share a high degree of identity with those identified from other luteoviruses. Among luteovirus coat protein sequences in general, several highly conserved domains were identified whereas other domains differentiate MAV-PS1 and PAV isolates from NY-RPV and other luteoviruses. Sequence similarities and differences among BYDV coat proteins (approx. 22K) are consistent with the serological relationships exhibited by these viruses. Amino acid sequence comparisons between BYDV isolates that share common aphid vectors indicate that it is unlikely that these coat proteins are involved in aphid specificity.

Introduction

Luteoviruses cause yellowing diseases on a wide range of host plants (Matthews, 1982). They are transmitted obligately by aphid vectors in a persistent, circulative manner and are limited to the phloem tissue of the host plant. Many luteoviruses exhibit some degree of serologically inter-relatedness (Waterhouse et al., 1988; Rochow & Duffus, 1981), and yet they can be grouped into clusters based on the closeness of the serological relationships.

Barley yellow dwarf virus (BYDV), the type member of the luteoviruses (Matthews, 1982), comprises a group of serologically related viruses that infect barley, oats, wheat, rice and other graminaceous hosts (Rochow, 1970a). Isolates of BYDV were originally distinguished by, and named according to, their predominant aphid vectors (Rochow, 1970a). They can also be separated into two major groups based on serological relationships (Aapola & Rochow, 1971; Rochow, 1970a; Rochow & Carmichael, 1979; Rochow & Duffus, 1981), cytopatho-

† Present address: USDA-ARS, BARC-West, Building 006, Beltsville, Maryland 20705, U.S.A.
‡ Present address: Department of Plant Science, University of Arizona, Tucson, Arizona 85721, U.S.A.

0000-9689 © 1990 SGM

logical ultrastructure of infected cells (Gill & Chung, 1979), and the profiles of dsRNAs obtained from infected tissue (Gildow et al., 1983). Group 1 includes the representative isolates: MAV, transmitted by Macrosiphum (= Sitobion) avenae Fabr.; PAV, transmitted by M. avenae and Rhopalosiphum padi L.; and SGV, transmitted by Schizaphis graminum Rond. (Rochow, 1970a). Group 2 includes the representative isolates: RMV, transmitted by R. maidis Fitch.; and RPV, transmitted by R. padi (Rochow, 1970a).

The differences in both serological and aphid vector relationships among BYDV isolates are presumed to reflect differences in viral coat proteins (Rochow, 1970b). The characterization of coat protein genes from different isolates should therefore provide a more detailed understanding of the serological and vectorial relationships exhibited by these viruses. Coat protein nucleotide sequences have been obtained from several luteoviruses, including potato leafroll virus (PLRV) (Kawchuk et al., 1989; Prill et al., 1989), beet western yellows virus (BWYV) (Veidt et al., 1988) and a PAV-like BYDV isolate from Victoria, Australia (Miller et al., 1988a), herein referred to as Vic-PAV. Furthermore, complete genome nucleotide sequences have been reported for BWYV (Veidt et al., 1988), PLRV (Mayo et
oligonucleotide sequencing primers were prepared (Laboratory for there-fore presumed to identify the viral coat protein-coding region. Generated by single and double enzyme restriction digests and by BYDV isolates were similarly prepared in 2gtl 1, pUC18 and pGEM-3Z (Promega). cDNA libraries representing the P-PAV and NY-RPV specific polyclonal antisera prepared against purified virus, and colony hybridizations to identify similar sequences in clones from the lab. 31-generated deletions in existing cDNA clones were identified (Putney et al., 1981); (iv) to sequence regions inaccessible by other methods, synthetic amino acid sequences of the coat proteins.

Methods

Viruses, cDNA cloning, and library screening. The BYDV isolates used were as follows: (i) a subculture of the New York MAV (NY-MAV) isolate of Rochow, maintained at Purdue University and designated MAV-PS1 (Barbara et al., 1987; Lister & Sward, 1988), (ii) the P-PAV isolate of Hammond (Hammond et al., 1983) and (iii) the New York RPV isolate of Rochow (NY-RPV) (Rochow, 1970a) as maintained at Purdue University. All virus isolates were maintained in oats (Avena sativa L. cv. Clintland-64) and checked by ELISA for cross-contamination as previously described (Barbara et al., 1987). Virus was purified and the RNA extracted as described for the MAV-PS1 isolate (Barbara et al., 1987).

The construction of the MAV-PS1 cDNA library, the cloning of the cDNA into agt1, and the subcloning from agt1 into pUC18 (Vieira & Messing, 1982) were previously described by Barbara et al. (1987). Subsequently, the MAV-PS1 cDNA was cloned into pGEM-3Z (Promega). cDNA libraries representing the P-PAV and NY-RPV BYDV isolates were similarly prepared in agt1, pUC18 and pGEM-3Z. The agt1 libraries were screened for immunologically recognizable lacZ fusion proteins (Young & Davis, 1983) with serotype-specific polyclonal antiserum prepared against purified virus, and therefore presumed to identify the viral coat protein-coding region. Immunoreactive 4 clones were then used as hybridization probes in colony hybridizations to identify similar sequences in clones from the plasmid libraries. Hybridization probes were labelled with [α-32P]dATP (Amersham) by either nick translation (Maniatis et al., 1982) (nick translation kit, Amersham) or random multiprimer labelling (Feinberg & Vogelstein, 1983) (multimprime labelling system, Amersham). Restriction maps representing each viral genome were generated by single and double enzyme restriction digestes and by Southern blot hybridization between different restriction fragments.

cDNA sequencing and verification of the NY-RPV coat protein gene. Plasmid DNA isolated by alkaline lysis techniques (Birnboim & Doly, 1979) was sequenced by the dideoxyribonucleotide chain termination method (Sanger et al., 1977) with a modified T7 DNA polymerase (Sequenase). Both DNA strands from clones representing the antigenic regions of the viral genomes and neighbouring regions were sequenced by one or more of the following strategies: (i) restriction fragments representing overlapping regions of the genome were subcloned; (ii) Bal 31-generated deletions in existing cDNA clones were identified (Guo et al., 1983); (iii) exonuclease III/mung bean nuclease-generated nested deletions were identified in existing cDNA clones (Putney et al., 1981); (iv) to sequence regions inaccessible by other methods, synthetic oligonucleotide sequencing primers were prepared (Laboratory for Macromolecular Structure, Department of Biochemistry, Purdue University). Sequence analyses were performed with Microgenie (Beckman) and the Genetics Computer Group sequence analysis software, version 6.1 (Devereux et al., 1984).

Results

Identification and verification of the coat protein-coding regions

Based on separation by SDS–PAGE, minor differences in the size of the capsid proteins of the MAV-PS1, P-PAV and NY-RPV BYDV isolates were detected (Fig. 1). With reference to M₄ markers, the migration of MAV-PS1, P-PAV and NY-RPV coat proteins corresponded to 21.8K, 21.9K and 22.9K respectively. In the course of analysing the nucleotide sequence of the MAV-PS1, P-PAV and NY-RPV genomes (P. P. Ueng et al. and J. R. Vincent et al., unpublished results), single open
Table 1. Comparison of luteovirus coat proteins and 17K internal ORF-encoded proteins

<table>
<thead>
<tr>
<th>Virus</th>
<th>Coat protein</th>
<th>Internal ORF-encoded protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M_t$</td>
<td>Bases</td>
</tr>
<tr>
<td>MAV-PS1</td>
<td>21.9K</td>
<td>597</td>
</tr>
<tr>
<td>P-PAV</td>
<td>22.0K</td>
<td>600</td>
</tr>
<tr>
<td>NY-RPV</td>
<td>22.2K</td>
<td>612</td>
</tr>
<tr>
<td>Vic-PAV</td>
<td>22.0K</td>
<td>600</td>
</tr>
<tr>
<td>BWYV (FL1)</td>
<td>22.5K</td>
<td>606</td>
</tr>
<tr>
<td>BWYV (GB1)</td>
<td>22.5K</td>
<td>606</td>
</tr>
<tr>
<td>PLRV-1</td>
<td>23.2K</td>
<td>624</td>
</tr>
<tr>
<td>PLRV-2</td>
<td>23.2K</td>
<td>624</td>
</tr>
</tbody>
</table>

reading frames (ORFs) were identified that corresponded to the apparent size of these capsid proteins. The coding region for the putative MAV-PS1 coat protein is 597 bp and would encode a 21934 protein (Fig. 2a). The coding region for the putative P-PAV coat protein is 600 bp and would encode a 21981 $M_t$ protein (Fig. 2b), whereas the NY-RPV putative coat protein-coding region is 612 bp and corresponds to a 22190 $M_t$ protein (Fig. 2c). All three regions of these cDNA sequences also contained a second ORF within the coat protein-coding region in a different frame (Fig. 2). These internal ORFs, which would produce proteins of $M_t$ 17181, 17124 and 17211 for MAV-PS1, P-PAV and NY-RPV, respectively, are thought to correspond to the VPg for each of these BYDV isolates (Miller et al., 1988a; Murphy et al., 1989).

Sequencing of the MAV-PS1 and P-PAV coat protein-coding regions also identified occasional nucleotide heterogeneities. These are located at position 300 for MAV-PS1 and positions 225 and 370 for P-PAV (Fig. 2). Although these variations generally did not alter the amino acid sequence (except P-PAV position 225), it is not known whether these differences represent authentic variability within the viral genome or whether these substitutions represent cloning artefacts.

The lengths of the identified coding regions, and the predicted molecular sizes of the capsid proteins and the VPgs correspond to similar proteins that have been characterized from other luteoviruses (Table 1). In vitro translation of the mRNA transcript from pRPVCP5 yielded three major proteins, the largest of which comigrated with the authentic NY-RPV coat protein (Fig. 3, lanes 2, 3 and 4); this provided additional evidence that these ORFs corresponded to the coat protein sequences. The 22K protein was also immunoprecipitated with NY-RPV antisera. Removal of the first AUG from pRPVCP5 yielded only smaller $M_t$ products (Fig. 3, lane 5). Based on the $M_t$ markers, both of these proteins migrated corresponding to $M_t$ values larger than expected for the VPg. Presumably one of these proteins corresponds to the internal ORF for the 17K protein; the other may represent a truncated in vitro translation product.

Sequence comparisons between coat protein-coding regions for BYDV and other luteoviruses

The amino acid sequence identity of the NY-RPV coat protein ORF with those of MAV-PS1 and P-PAV, as well as the similarity of these proteins with coat proteins of other luteoviruses (Kawchuk et al., 1989; Miller et al., 1988a; Prill et al., 1989; Veidt et al., 1988), provide convincing evidence for their identification. When the nucleotide and amino acid sequences of these luteovirus coat proteins were gapped and aligned to maximize their similarities, obvious relationships between them were apparent. At the nucleotide level, identities extend from near 55%, in comparisons between NY-RPV (or BWYV or PLRV) and the MAV-PS1 or Vic-PAV isolates, to 95% when the different PAV isolates are compared to each other. Specifically, comparisons of MAV-PS1 and NY-RPV with P-PAV show 76% and 56% identity, respectively (Table 2). Most noticeable in these comparisons is that the NY-RPV sequence has greater identity to the sequences from BWYV and PLRV than to those from the other BYDV isolates.

Coat proteins of the various luteoviruses can be separated into two domains: an amino-terminal domain of variable length which ranges from 60 to 69 amino acids (Fig. 4), and a carboxy-terminal (C-terminal) domain which essentially contains an identical number of amino acids for all luteovirus coat protein sequences (Fig. 5). A distinctive feature of the amino-terminal domain of the luteovirus coat protein is the occurrence of two, highly basic regions separated by a region of non-
(a) BYDV coat protein gene: MAV-PS1 isolate

```
MNSVGRNHRNRNMRNPGRARRRVRVSAVRMVVVQPNRA
MAQGEQGALAQFGELWLSNPIEP
UGAAAUUCAGUACGCGGAGAAGAACACCAAGGAGGCGAGAGCCGAGGAGAUGAGGCGGAGGAGAUGAGGAGGAGCAGGCAGGAGGCAGGAGGAGAUGAGG
10 20 30 40 50 60 70 80 90 100
GPKRRARRRRRGGGAGNLISGPAGRTTEVFVFVSVNDLDQMDVELVDAAQEEEGGILYLIDQAGLRYSYSQSTTL
GACCAGAAACCGAGACGACGACGACGACGACGACGACGACGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAGGG
110 120 130 140 150 160 170 180 190 200 210

KANSSTIKFGPDLSQCPALSAGGILKSYLKYKITN
RPTPQGQSSSSVPTFRNAQRFQVEYSSPTTFTSQT
AGGCGGUAUCUCAGGCAAUCAUGUAGUAGUACGCAGCCUAUCACAGCAUCAGCAUCAGCAUCAGCAUCAGCAUCAGCAUCAGCAUCAGCAUCAGCAUCAG
220 230 240 250 260 270 280 290 300 310

GKKEFRESTVNFYMLYK
HSPSQNQPKPSPPNRLTGRNSGRVRU
UCUAUCAUCUAUCUAUAAACGACCACCCACACGGCAGCGCCGCGACUGCAGCAAGCCGAGGCGGAGGCGGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGG
320 330 340 350 360 370 380 390 400 410 420
ANGSTSĐTQAGFIITIRVANMTPKU
GGGAAUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
530 540 550 560 570 580 590 600
```

(b) BYDV coat protein gene: P-PAV isolate

```
MNSVGRGPRRANQNQTRRRRTTRPVVVVQPNRM
MAQEGGAVEQFCQLWLSNPIEP
UGAAAUUCAGUAGCGGAGAGCAGGAAGCAGGAGGCGAAGAAACAAAGGCAAAACGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG
10 20 30 40 50 60 70 80 90 100
AGPRRMRGRRKGGGANPFVRPFRTGGTEVFVFVFSDN
QDPDDEMVDAREESEQILYLDQAGLRYSYSQSTT
CGAGCCCGACGCGCGAGGAAAGGCGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG
110 120 130 140 150 160 170 180 190 200 210

LKANSSTIKFGPSQLCPSALSDGIILKSYHRYKULTKPTTPGQSNASPVYRWAORFQOTEYLSPSTTVTSSQ
CUJUAAACGCAACUCUCGGCGGCAAUCAAAUACGGCGCAGCAGCAAGCCGCGGCAAGCAGCAAGCCGCGGCAAGCAGCAAGCCGCGGCAAGCAGCAAGCCGCGG
220 230 240 250 260 270 280 290 300 310

SIRVFKSHASATTAGAIFIELDTACKQSAINGSYIVSVLSSHSITRPOLRPAALSSNSSPANQDPWVATL
AGUAUCCCGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
320 330 340 350 360 370 380 390 400 410 420
```
Fig. 2. Nucleotide sequences for the coat protein genes of MAV-PS1, P-PAV and NY-RPV isolates of BYDV. The deduced amino acid sequence for the coat protein is shown on the top line for each isolate. The internal ORF-encoded sequence corresponding to a 17K protein is shown on the second line. Alternative nucleotides determined from the sequencing of multiple clones representing the coat protein region are indicated in parentheses below the nucleotide sequence. The corresponding alternative amino acids are indicated in parentheses above the amino acid sequence. (a) MAV-PS1; (b) P-PAV; (c) NY-RPV.

polar amino acids (Fig. 4). The amino acids found within the two basic regions are nearly all arginine residues and, as a result, arginine constitutes 60 to 70% of the amino acids found within the amino-terminal domain.

The C-terminal domains from luteovirus coat proteins show a high degree of amino acid conservation. Thirty-five percent of the amino acids in the C-terminal domains are positioned identically (Fig. 5). Many of these amino acids are clustered together forming highly conserved regions. Multiple sequence analyses (Devereux et al., 1984) of the C-terminal amino acid sequences from all of these luteoviruses identified a consensus sequence, in addition to separate consensus sequences for the MAV-PS1/PAV group and the NY-RPV/PLRV/BWYV group. These analyses showed that whereas some regions are conserved among all luteo-
viruses, other regions divide them into a group containing BYDV isolates MAV-PS1, Vic-PAV and P-PAV and a group containing BYDV NY-RPV, BWYV and PLRV.

**Discussion**

The viral coat protein sequences for the BYDV isolates MAV-PS1, P-PAV and NY-RPV share a high degree of homology. However, nucleotide and amino acid sequences of NY-RPV are more similar to those of BWYV and PLRV than to other BYDV isolates. The degree of sequence similarity among these coat protein sequences is consistent with the serological relationships by which BYDV isolates have been grouped. Thus the group 1 isolates MAV-PS1 and P-PAV share 76% nucleotide and 71% amino acid identity and, depending on the antisera and serological test procedures used, can show varying degrees of serological cross-reactivity. The P-PAV and Vic-PAV (Miller *et al.*, 1988a) isolates share 95% nucleotide and 97% amino acid identity and are of the same serotype. The distantly related group 2 isolate NY-RPV shares only 56% nucleotide and 50% amino acid identity with the P-PAV isolate and has long been regarded as serologically distinct. Consistent with these observations is the high degree of sequence identity between NY-RPV and BWYV, which have been shown to be serologically related (Rochow & Duffus, 1978, 1981).

The amino-terminal sequences (roughly 60 amino acids) of BYDV, BWYV and PLRV contain a high proportion of basic amino acids clustered into two regions. They are composed predominantly of arginine residues and are separated by relatively non-polar amino acids. The conservation of this organization among luteoviruses and the presence of two highly basic regions suggest that these features may be involved in structural protein–RNA interactions within the viral capsid (Harrison, 1983).

**Table 2. Nucleotide and deduced amino acid sequence similarity between the MAV-PS1, P-PAV, NY-RPV isolates of BYDV and other luteovirus 22K coat protein-coding regions**

<table>
<thead>
<tr>
<th>Nucleotide similarity (%)</th>
<th>MAV-PS1</th>
<th>NY-RPV</th>
<th>Vic-PAV</th>
<th>BWYV (FL1)</th>
<th>BWYV (GB1)</th>
<th>PLRV-1</th>
<th>PLRV-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-PAV</td>
<td>76-6</td>
<td>54-8</td>
<td>76-1</td>
<td>57-0</td>
<td>57-8</td>
<td>53-0</td>
<td>53-1</td>
</tr>
<tr>
<td>P-PAV</td>
<td>56-4</td>
<td>95-0</td>
<td>57-4</td>
<td>56-3</td>
<td>57-6</td>
<td>57-6</td>
<td></td>
</tr>
<tr>
<td>NY-RPV</td>
<td>55-2</td>
<td>65-4</td>
<td>65-3</td>
<td>69-4</td>
<td>69-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Deduced amino acid similarity (%)</th>
<th>MAV-PS1</th>
<th>NY-RPV</th>
<th>Vic-PAV</th>
<th>BWYV (FL1)</th>
<th>BWYV (GB1)</th>
<th>PLRV-1</th>
<th>PLRV-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-PAV</td>
<td>71-3</td>
<td>47-8</td>
<td>71-8</td>
<td>48-1</td>
<td>49-0</td>
<td>46-9</td>
<td>47-4</td>
</tr>
<tr>
<td>P-PAV</td>
<td>50-7</td>
<td>97-0</td>
<td>48-1</td>
<td>47-4</td>
<td>47-4</td>
<td>47-8</td>
<td></td>
</tr>
<tr>
<td>NY-RPV</td>
<td>48-8</td>
<td>66-2</td>
<td>65-2</td>
<td>65-6</td>
<td>66-2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Reciprocal comparisons between isolates and self-comparisons are not presented and are indicated by (–). References for other luteovirus coat protein sequences are: Vic-PAV, Miller *et al.* (1988a); BWYV (FL1) and BWYV (GB1), Veidt *et al.* (1988); PLRV-1, Prill *et al.* (1989); PLRV-2, Kawchuk *et al.* (1989).
### BYDV coat protein sequences

<table>
<thead>
<tr>
<th>Luteovirus</th>
<th>MAV-PS1</th>
<th>P-PAV</th>
<th>Vic-PAV</th>
<th>NY-RPV</th>
<th>PLRV-1</th>
<th>PLRV-2</th>
<th>BWYV (FLI)</th>
<th>BWYV (GBI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAV-PS1</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>P-PAV</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>Vic-PAV</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>NY-RPV</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>PLRV-1</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>PLRV-2</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>BWYV (FLI)</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
<tr>
<td>BWYV (GBI)</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
<td>PNNSVGRR</td>
<td>MN NSVGRR</td>
<td>MNNSVGRR</td>
<td>MNNSVGRR</td>
<td>MNSTNVQRTNG</td>
<td>MNSTNVQRTNG</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison of the amino acid sequences from amino-terminal regions of luteovirus coat proteins. The amino acid sequences of the luteovirus coat proteins were aligned to maximize their relatedness. When similar isolates are presented, only the amino acid differences between those isolates are shown; identical amino acids are represented by dots. Regions high in basic amino acid content are represented as boxed areas, and the basic amino acid residues are underlined. References: PLRV-1, Prill et al. (1989); PLRV-2, Kawchuk et al. (1989); BWYV, Veidt et al. (1988).

### Fig. 5. Comparison of the amino acid sequences from carboxy-terminal regions of luteovirus coat proteins. BWYV and the NY-RPV isolate of BYDV contain 138 amino acids in this region whereas PLRV and the PAV and MAV-PS1 isolates of BYDV contain 139 amino acids. When similar isolates are presented, only the amino acid differences between those isolates are shown; identical amino acids are represented by dots. Conserved regions representing consensus sequences identified from multiple sequence analyses (Devereux et al., 1984) are indicated by boxed areas. Asterisks above the amino acid sequence indicate that an amino acid is identically located in the carboxy-terminal region of all the luteoviruses. References: PLRV-1, Prill et al. (1989); PLRV-2, Kawchuk et al. (1989); BWYV, Veidt et al. (1988).

Although similarity comparisons among BYDV and other luteovirus coat protein sequences may provide an explanation for serological differences, the similarity in the C-terminal sequences of these proteins suggests that they all might show some degree of serological inter-relatedness. In this regard, polyclonal and monoclonal antisera to MAV isolates, and to P-PAV and NY-RPV have been shown to recognize both homologous and heterologous BYDV isolates in indirect ELISA (Diaco et al., 1986a) and in serologically specific electron microscopy (Diaco et al., 1986b). Indirect ELISA techniques also indicated serological cross-reactivity among a BYDV isolate ('BYDV-PAV-IL'), a BWYV isolate ('BWYV-CA') and an isolate of soybean dwarf virus, another luteovirus (Hewings & D'Arcy, 1986). Because BYDV dissociates in the alkaline buffers used for
antigen coating (Diaco et al., 1986a), these indirect ELISA procedures probably analysed the serological relationships of dissociated virus particles.

Because of the high degree of amino acid sequence similarity between luteovirus coat proteins, even serotype-specific antisera may thus also identify common epitopes and could, in fact, be used as general probes for luteoviruses, providing virus preparations are appropriately treated (e.g. dissociated). Immunological detection procedures which tend to preserve the structural integrity of luteoviruses, such as double antibody sandwich ELISA, emphasize isolate-specific detection (Rochow & Carmichael, 1979), but procedures involving disrupted virus, including some indirect ELISAs or Western blot analyses, may permit immunological cross-reactions detecting many luteoviruses, and indicating a broader spectrum of serological relationships.

Often, BYDV isolates are most efficiently transmitted by specific aphid vectors. Such specificity is a property of specific interactions of the viral capsid with receptor sites on the accessory salivary gland membranes of the vector (Gildow & Rochow, 1980; Gildow, 1987). Of the isolates used in the present study, the P-PAV isolate and other PAV isolates are non-specifically transmitted by both R. padi and M. avenae (Rochow, 1970a). Thus R. padi can transmit both P-PAV and NY-RPV, and M. avenae can transmit both P-PAV and MAV-PS1. NY-RPV shares close sequence similarity with both BWYV and PLRV, but is efficiently transmitted by R. padi whereas BWYV and PLRV are both efficiently transmitted by Myzus persicae. These vector relationships allow evaluation of the coat protein sequences for regions likely to be involved in aphid specificity. Thus amino acids which are similarly located in NY-RPV and BWYV or PLRV are unlikely to be specifically involved in the aphid transmissibility of BYDV. Similarly, those amino acids similarly located in P-PAV and MAV-PS1 are unlikely to be specifically involved in the transmissibility of NY-RPV and P-PAV by R. padi. Elimination from consideration of the above amino acid pairs then allows us to identify amino acid pairs which are unique to PAVs and NY-RPV, and thus sequences potentially involved in R. padi transmission of PAVs and NY-RPV.

Analyses of deduced amino acid sequences of the C-terminal regions of luteovirus coat proteins reveal that there is only one pairing of amino acids which is unique to both P-PAV and NY-RPV, alanine at P-PAV position 120 and NY-RPV position 125 (Fig. 5). The other luteoviruses have different amino acids at this location, but these differences result from single base changes from the GCC codon found in both P-PAV and NY-RPV. Furthermore, both BWYV and PLRV have the related amino acid serine at this position. At all other positions within the P-PAV or NY-RPV C-terminal domains, either P-PAV and NY-RPV have different amino acids, or an identical amino acid can be found in one of the other luteovirus sequences. In the light of these observations, it seems unlikely that the C-terminal region of luteovirus coat proteins is responsible for vector specificity. In icosahedral RNA viruses, the N-terminal regions of the coat proteins are typically located within the interior of the virus particle (Rossmann & Johnson, 1989), and thus this region is also unlikely to be involved in aphid specificity. Therefore, it is likely that another protein on the viral capsid is involved. In this regard, it has been proposed that a second structural protein which has been identified in Vic-PAV virions is involved in aphid specificity (Waterhouse et al., 1989). In recent work in our laboratory (P. P. Ueng et al. and J. R. Vincent et al., unpublished results), similar proteins have been identified by Western blot analyses, the sizes of which correspond to ORFs identified by sequencing of cDNA produced from the three BYDV isolates discussed herein.

We thank C. Logan and H. Laffoon for their photographic services. This work was funded by Grant 88-37263-3855 from the Competitive Research Grants Office of the U.S. Department of Agriculture and Grant 1484670 from The Quaker Oats Company. The Genetics Computer Group Sequence Analysis Software Package for the VAX was supported by NIH grant No. AI27713. The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X17260 for BYDV MAV-PS1, X17261 for BYDV P-PAV and X17259 for BYDV NY-RPV. This is Journal Series Article 12326 of the Purdue University Agricultural Experiment Station.

References


(Received 21 May 1990; Accepted 23 August 1990)