Evidence for similarity-assisted recombination and predicted stem–loop structure determinant in potato virus X RNA recombination

Heidrun-Katharina Draghici1 and Mark Varrelmann1,2

Correspondence
Mark Varrelmann
varrelmann@ifz-goettingen.de

1Department of Crop Sciences, Section Plant Virology, University of Göttingen, Grisebachstrasse 6, D-37077 Göttingen, Germany
2Institute of Sugar Beet Research, Department of Phytopathology, Holtenser Landstrasse 77, D-37079 Göttingen, Germany

Virus RNA recombination, one of the main factors for genetic variability and evolution, is thought to be based on different mechanisms. Here, the recently described in vivo potato virus X (PVX) recombination assay [Draghici, H.-K. & Varrelmann, M. (2009). J Virol 83, 7761–7769] was applied to characterize structural parameters of recombination. The assay uses an Agrobacterium-mediated expression system incorporating a PVX green fluorescent protein (GFP)-labelled full-length clone. The clone contains a partial coat protein (CP) deletion that causes defectiveness in cell-to-cell movement, together with a functional CP+3'-ntr transcript, in Nicotiana benthamiana leaf tissue. The structural parameters assessed were the length of sequence overlap, the distance between mutations and the degree of sequence similarity. The effects on the observed frequency of reconstitution and the composition of the recombination products were characterized. Application of four different type X intact PVX CP genes with variable composition allowed the estimation of the junction sites of precise homologous recombination. Although one template switch would have been sufficient for functional reconstitution, between one and seven template switches were observed. Use of PVX–GFP mutants with CP deletions of variable length resulted in a linear decrease of the reconstitution frequency. The critical length observed for homologous recombination was 20–50 nt. Reduction of the reconstitution frequency was obtained when a phylogenetically distant PVX type Bi CP gene was used. Finally, the prediction of CP and 3'-ntr RNA secondary structure demonstrated that recombination-junction sites were located mainly in regions of stem–loop structures, allowing the recombination observed to be categorized as similarity-assisted.

INTRODUCTION

RNA viruses are extremely flexible living entities, adapting quickly to changing environmental conditions (Domingo & Holland, 1997; Roossinck, 1997), and combine high population densities with small genomes and template-replicating enzymes. Viral replicases possess a high mutation rate, due to missing proofreading abilities of the RNA-dependent RNA polymerases (RdRps) (Domingo et al., 1999, 2002). The mechanism of RNA recombination, which joins two molecules of different parental/replicational origin, has two main functions. In addition to representing a source of variation, recombination can rescue viral templates by reversing deleterious mutations (Aaziz & Tepfer, 1999a; Fernández-Cuartero et al., 1994; White & Morris, 1994).

Despite its important function for variability and evolution of RNA viruses, knowledge about the underlying mechanism of recombination, including the identification and functionality of the RdRp domains involved, is still uncertain. Knowledge of RNA primary (sequence composition and similarity) and secondary structure parameters, which can modify recombination, is available for only a few species. Evidence for the occurrence of the RNA recombination phenomenon either by means of in silico sequence comparison (reviewed by Chare & Holmes, 2006) or by its occurrence in mixed infections (Aaziz & Tepfer, 1999a; de Wispelaere et al., 2005; Meier & Truve, 2006), or under experimental conditions by applying high selection pressure, has been given for numerous plant virus species (Aaziz & Tepfer, 1999b; Simon & Bujarski, 1994). However, knowledge concerning RNA recombination mechanisms and parameters affecting these processes has been obtained mainly from a limited number of positive-strand RNA model virus species, namely brome mosaic...
virus (BMV) and cowpea chlorotic mottle virus, both belonging to the genus Bromovirus (Aleksa et al., 2005; Allison et al., 1996; Bujarski et al., 1998; Figlerowicz & Bujarski, 1998) in supergroup III (van der Heijden & Bol, 2002). Supergroup II RNA recombination model viruses are tomato bushy stunt virus (TBSV) (Panavas & Nagy, 2003), turnip crinkle virus (TCV) (genus Carnovirus) (Cascone et al., 1993; Zhang et al., 1991) and the RNA bacteriophage Qβ (Brown & Gold, 1996). Animal-infecting recombination model viruses include poliovirus (Jarvis & Kirkegaard, 1992) and mouse hepatitis virus (Banner et al., 1990; Fu & Baric, 1992).

It is generally accepted that replicative RNA recombination is driven by the ability of the viral replicase to switch templates during nascent-strand synthesis from a donor to an acceptor RNA molecule and by the use of nascent strands as primers to complete synthesis on the acceptor strand. Nevertheless, the involvement of host factors cannot be excluded. Several variants of nascent-strand transfer are conceivable (Gmyl & Agol, 2005; Nagy & Simon, 1997), but experimental evidence is still lacking. Important structural factors of the template or nascent strand that are used for classification have been identified in these model viruses and can be subdivided into primary factors, such as sequence similarity, AU-rich motifs, sequence complementarity and replicative cis elements, and secondary structures, such as stem–loops (Bruyere et al., 2000; Nagy & Bujarski, 1993, 1995; Nagy et al., 1998; Olsthoorn et al., 2002; Shapka & Nagy, 2004). So far, non-replicative RNA recombination, possibly based on RNA breakage and ligation, has only been detected in bacteriophage Qβ and polioviruses (Chetverin et al., 1997; Chetverina et al., 1999; Gallei et al., 2004; Gmyl et al., 1999, 2003), but not yet in plant viruses.

Recombination model viruses that belong to the genera Bromovirus and Tombusvirus possess icosahedral particle morphology. Tombusviruses encode one molecule of replication-associated protein and a readthrough polymerase, both necessary for replication. In contrast, bromoviruses possess a replicase molecule with polymerase and helicase functions plus methyltransferase on two separate proteins (reviewed by van der Heijden & Bol, 2002). We have recently established a recombination assay for potato virus X (PVX) (Draghici & Varrelmann, 2009), another plus-strand RNA virus with flexuous, rod-shaped particle morphology, representing a replication model virus with a single replicase molecule (Batten et al., 2003; Verchot-Lubicz et al., 2007). This in vivo assay is based on Agrobacterium tumefaciens-mediated expression of a green fluorescent protein (GFP)-labelled PVX isolate with a partial coat protein (CP) open reading frame (ORF) deletion in parallel with an intact CP transcript plus 3′ non-translated region (ntr). When we applied PVX replicase insertion-scanning mutants (Draghici et al., 2009), we were able to identify the involvement of a separate viral replicase domain in replicational RNA recombination, thereby supporting previous mutational studies of BMV (Dzianott et al., 2001; Figlerowicz et al., 1997, 1998; Nagy et al., 1995) and cucumber necrosis virus (Panaviene & Nagy, 2003) replicase.

Applying this in vivo recombination assay, which exerts a high selection pressure on the generation of recombinants, we aimed to study efficiently the parameters affecting PVX RNA recombination. PVX CP transcripts from isolates harbouring several mutations were used to determine the number and location of recombination-junction sites. Several PVX mutants containing CP deletions with different lengths of overlap of the intact CP gene were tested for their influence on the observed time point and frequency of reconstitution. Another CP gene derived from a different PVX type with only limited sequence similarity was tested for possible occurrence of non-homologous recombination events, frequency of reconstitution and composition of movement-competent recombinant products. By using CP and 3′-ntr secondary-structure predictions, we aimed to find evidence for the influence of RNA secondary structure on the location of recombination-junction sites.

**RESULTS**

The relative frequency of functional PVX CP gene reconstitution by means of recombination is proportional to the length of homologous sequence overlap

When PVX-GFP-ΔCP transcripts were co-expressed with CP–3′-ntr transcripts in *Nicotiniana benthamiana* leaf patches, the agroinfiltration of 65 plants (each inoculated four times) resulted in 60 systemically PVX–GFP-infected plants at 14 days post-infection (p.i.) (Draghici & Varrelmann, 2009). Such a frequent occurrence of recombination, leading to the reconstitution of a complete CP (711 nt in length), was possible because two continuous homologous stretches of CP, nt 1–514 and 646–711, plus the entire 3′-ntr of 72 nt in length, were available for template switching. The intact CP + 3′-ntr transcript, however, did not contain any nucleotide mismatch compared with the deletion mutant and therefore did not allow determination of the RdRp template switch for subsequent reconstitution. We intended to analyse the relationship between frequency of CP gene reconstitution by means of recombination and the length of the overlapping sequence. This was achieved by the combined use of PVX CP mutants carrying deletions of variable length plus intact CP-UK3 transcripts. We expected a proportional relationship and a possible minimal length of sequence overlap to be necessary for successful CP gene reconstitution. Moreover, we aimed to find evidence for whether PVX recombines mainly via replicative RdRp template switching between homologous sequences or whether RNA-ligation activity might represent an alternative recombination mechanism. Several 35S-PVX-GFP-CP deletion mutants were constructed (35S-PVX-GFP-CP-0, -10, -20, -50, -101, -200, -300, -401, -500 and -600 in Fig. 1) and each was co-expressed with 35S-CP-UK3 in 20
infiltrated leaf patches in two independent experiments. Virally expressed GFP fluorescence was monitored over time by using UV illumination, from 3 until 17 days p.i. (Table 1).

Spread of recombinant virus was first detected at 8 days p.i., when constructs 35S-PVX-GFP-CP-600 + 35S-CP-UK3 were used for co-infiltration. When applying construct 35S-PVX-GFP-CP-500 to -CP-101 + 35S-CP-UK3, recombinant virus movement was first observed at 10 days p.i. When testing 35S-PVX-GFP-CP-50, virus-encoded fluorescence was not observed before 12 days p.i. In contrast, there was no spread of PVX during the time period analysed when mutants -CP-20, -CP-10 and -CP-0 + 35S-CP-UK3 were applied (Table 1). In addition to this observation, the total number of agroinfiltrated patches displaying virus movement observed at 17 days p.i. correlated positively with the length of sequence overlap between deletion mutant and intact CP transcript.

**Use of intact PVX CP genes with sequence variability enables determination of the composition of the recombinant CP sequence and the number of RdRp template switches**

The recombination assay did not allow determination of the number of template switches or preferred recombination sites up- and downstream (3' and 5') of the CP gene (nt 514–646) deletion because the overlapping parental sequences were identical. To apply intact genes possessing nucleotide exchanges that can serve as markers for this

### Table 1. Time-dependent occurrence of PVX recombination in relation to the length of sequence overlap available for recombination

Values shown are the number of *N. benthamiana* leaf patches infiltrated with 35S-PVX-CP displaying CP deletions of different length (20 patches each treatment) and intact 35S-CP-UK3. I and II are independent repetitions of each experiment.

<table>
<thead>
<tr>
<th>35S-PVX-GFP-CP mutant</th>
<th>Time p.i. (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>CP-0 del 652 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-10 del 642 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-20 del 632 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-50 del 602 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-101 del 551 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-200 del 452 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-300 del 352 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-401 del 251 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-500 del 152 bp</td>
<td>0</td>
</tr>
<tr>
<td>CP-600 del 52 bp</td>
<td>1</td>
</tr>
</tbody>
</table>

---

Fig. 1. Schematic representation of 35S-PVX-GFP and CP deletion mutants derived thereof (not drawn to scale). Abbreviations: LB/RB, left and right T-DNA border sequences; 35S, 35S promoter of cauliflower mosaic virus (CaMV); RdRp 166, RNA-dependent RNA polymerase ORF; P25, P8 and P12, triple-gene-block ORFs; sgPr, subgenomic CP promoter; GFP, green fluorescent protein ORF; CP, coat protein ORF; 3'-ntr, 3' non-translated region; pA, CaMV poly(A) terminator; tnos, nopaline synthase transcriptional terminator; Pvull/Xhol, restriction sites. The complete PVX CP ORF includes 711 nt and is illustrated as a light grey box. CP deletion mutants 35S-PVX-GFP-CP-0, -10, -20, -50, -101, -200, -300, -401, -500 and -600 are marked by dashed lines.
purpose, four different PVX type X isolates were used for cloning and sequencing of the CP gene + 3′-ntr region (PVX-14, -17, -18 and -847). CLUSTAL_X (Thompson et al., 1997) alignment revealed 100% sequence identity of the 3′-ntr between all isolates (data not shown) and 20–31 nt exchanges in the entire 711 nt CP ORF (95.6–97.2% sequence similarity) compared with CP-UK3 (Fig. 2). Most of these mutations were silent and located at the 5′ end of the deletion in CP-UK3, leading to only three amino acid exchanges (data not shown). Therefore, we assumed functionality of chimeric CPs, and no selection of recombinants with specific compositions was expected. All 35S-driven CP + 3′-ntr expression clones (35S-CP-PVX-14, -17, -18 and -847) plus the 35S-PVX-GFP ΔCP control were applied in the assay. The treatment was tested in five N. benthamiana plants, each being inoculated at four positions. Surprisingly, in both independent repetitions, the first recombinants occurred at least 1 day earlier (at 7 instead of 8–9 days p.i.) when the divergent CP genes were applied instead of the CP-UK3 transcript (Table 2). However, the reconstitution frequency within the time period remained comparable.

To determine the sequence and variability of recombinants among different treatments and repetitions, five N. benthamiana plants from each treatment (at 10 days p.i.), displaying GFP fluorescence in non-inoculated leaves, were used for upper-leaf sampling, cloning and sequencing of recombinant CP genes. In principle, the sequence composition of the recombinant CP genes does not enable determination of recombinant generation via plus- or minus-strand synthesis, even if mutations in the parental sequences permit the identification of regions of possible template switching. Only for reasons of consistency and clear evaluation was template switching during RdRp minus-strand synthesis anticipated for all recombinants analysed. The 3′-ntr displayed 100% identity between all isolates and made it impossible to determine on which template the RdRp initiated minus-strand synthesis. As one template switch would be sufficient to reconstitute a functional CP if the initiation of transcription took place on the intact CP transcript, this was anticipated for all recombinants analysed. Sequence analysis of the 20 recombinants supported this hypothesis, showing that the

Fig. 2. CLUSTAL_X alignment of the CP ORFs of different PVX type X isolates and PVX-UK3. Asterisks indicate PvuII CP gene deletion in the UK3 template produced by 35S-PVX-GFP ΔCP. Nucleotide numbering is according to the PVX-UK3, -14, -17, -18 and -847 CP ORFs, respectively.
nucleotide at the most 3′-end position (nt 702) in the CP ORF, divergent in all four different isolates, was different in 18 of 20 recombinants analysed.

Recombinant sequence alignment of the different treatments and repetitions revealed several interesting features. Independently of the four CP sequences, many recombinants from independent plants displayed variable properties [number of template switches (recombination events), mutations], which in some cases resulted in recombinants with variable amino acid composition (see Supplementary Fig. S1, available in JGV Online; summarized in Table 2). Although the deletion in the CP-UK3 gene could have been repaired by an RdRp template switch 5′-proximal to the PvuI deletion from the 35S-CP transcript to the intact PVX-UK3 template strand and completion of viral minus-strand synthesis, this did not occur frequently. Thirteen of 20 recombinants were a product of more than one template switch. In contrast, three recombinants with seven template switches (PVX-14 Rec4 and PVX-847 Rec1 and 2) were observed. Additionally, sequence analysis revealed that most of the recombinant sequences (14 of 20) were unique. No evidence for the selection of a specific region in the variable amino acids was found (Table 2). All amino acid differences between PVX-UK3 and PVX-14, -17, -18 and -847 were located in the N-terminal 60 aa (data not shown). The CP-UK3 deletion (aa 17–215) did not overlap with these mutations. Theoretically, this would allow the reconstitution of a recombinant CP with wild-type composition. However, several recombinants encoding chimeric or even the entire intact divergent CP sequence were detected (Table 2). Comparison between recombinants derived from one treatment, as well as between different treatments, did not reveal preferred regions for RdRp template switches or indications of recombinational hot spots. Even the recombination events leading to CP gene reconstitution by only a single template switch, which was found in seven recombinants, seemed to be distributed homogeneously across the CP ORF (Supplementary Fig. S1).

### Table 2. Properties of chimeric PVX recombinant CP genes derived from recombination of PVX-UK3 CP gene and divergent CP genes of four different PVX type X isolates

| PVX isolate* | Sequence divergence† | Infiltrated patches‡ | Recombinant no. analysed | Template switches§ | Amino acid composition selected | Recombinant properties|| |
|--------------|----------------------|---------------------|-------------------------|-------------------|-------------------------------|-------------------|
| PVX-14       | 24                   | P10T, I11T, S60N    | 16/20                   | 1                 | CP-UK3                        | 5                 |
|              |                      |                     |                         | 2                 | CP-UK3                        | 3                 |
|              |                      |                     |                         | 3                 | CP-UK3                        | 2                 |
|              |                      |                     |                         | 4                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 5                 | CP-UK3                        | 1                 |
| PVX-17       | 29                   | P10T, S42N, S60R    | 16/20                   | 1                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 2                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 3                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 4                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 5                 | Chimera (PNR)§§§              | Unique            |
| PVX-18       | 31                   | P10T, S42N, S60R    | 17/20                   | 1                 | Chimera (PSR)§§§              | 2                 |
|              |                      |                     |                         | 2                 | Chimera (PSR)§§§              | 1                 |
|              |                      |                     |                         | 3                 | Chimera (PSR)§§§              | Unique            |
|              |                      |                     |                         | 4                 | Chimera (PSR)§§§              | Unique            |
|              |                      |                     |                         | 5                 | UK3                           | Unique            |
| PVX-847      | 20                   | P10A, I11T, A24T    | 10/20                   | 1                 | Chimera (PTT)§§§              | Unique            |
|              |                      |                     |                         | 2                 | CP-847                        | Unique            |
|              |                      |                     |                         | 3                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 4                 | CP-UK3                        | Unique            |
|              |                      |                     |                         | 5                 | Chimera (PTT)§§§              | Unique            |

*PVX isolate used for CP transcript generation.
†Sequence divergence compared with CP-UK3 [no. of nucleotide exchanges (nt) or amino acid exchanges (aa)].
‡No. of infiltrated patches displaying PVX movement.
§No. of template switches detected.
||Unique or identical to recombinant no.
§§§Composition of the variable amino acids.

A CP transcript plus 3′-ntr from a phylogenetically distant PVX type Bi isolate reconstitutes a PVX-UK3 CP deletion mutant and enables a more precise determination of recombination sites

The long distance between nucleotide exchanges when different intact, type X CP genes (PVX-14, -17, -18 and
fluorescence (named PVX-4-Rec1 to -Rec15) were used for individual recombinants, 15 plants showing systemic GFP To determine the sequence composition of different that were tested in parallel. p.i., compared with 72 of 104 patches of CP-UK3 control infiltrated patches displayed virus movement at 14 days recombination frequency observed with CP-PVX-4 tran- with recombination with type X CP genes, the relative exchanges in the 3\textsuperscript{9}-ntr displaying only 75.1\% CP gene and 76.8\% CP gene + 3\textsuperscript{-ntr} sequence similarity to CP-UK3. As described by Santa Cruz & Baulcombe (1995), the CP-CPVX-4 Bi type isolate contained an insertion of 33 nt in the 3\textsuperscript{'} end (see Supplementary Fig. S2, available in JGV Online). For precise determination of the recombination- junction sites, the relatively uniform spread of 187 nucleotide exchanges in the CP ORF and five nucleotide exchanges in the 3\textsuperscript{-ntr} appeared promising. Compared with recombination with type X CP genes, the relative recombination frequency observed with CP-PVX-4 trans- scripts was reduced severely. Only 21 of 208 double- infiltrated patches displayed virus movement at 14 days p.i., compared with 72 of 104 patches of CP-UK3 control that were tested in parallel.

To determine the sequence composition of different individual recombinants, 15 plants showing systemic GFP fluorescence (named PVX-4-Rec1 to -Rec15) were used for cloning and sequencing of the CP gene + 3\textsuperscript{-ntr}. A CLUSTAL_X alignment of the CP gene plus 3\textsuperscript{-ntr} of PVX-UK3, PVX-4 and 15 recombinants revealed that all recombinants were unique at the nucleotide level (Supplementary Fig. S2). Again, a single template switch would have been sufficient to reconstitute a complete, functional CP gene. However, this was only detected in PVX-4-Rec1, 5, 7, 12 and 14. The remaining 10 recombinants exhibited between two and seven template switches. The 3\textsuperscript{-ntr} composition of the different recombi- nants showed that the most 3\textsuperscript{'} nucleotide exchange, at position 750 in the UK3 sequence (corresponding to position 783 in PVX-4; Santa Cruz & Baulcombe, 1995), T (PVX-UK3) \textrightarrow C (PVX-4), was shown to be PVX-4 in 10 of 15 recombinants (Supplementary Fig. S2). This demon- strated that, in the majority of the recombinations observed, the initiation of minus-strand synthesis had occurred on the intact PVX-4 transcript. In 15 recombin- ant CP genes analysed, 45 recombination-junction sites were observed, of which 27 were unique (Supplementary Fig. S2, sequences between shaded sections). CLUSTAL_X alignment of the amino acid composition revealed that only PVX-4-Rec2, 4, 10, 15 and 1, 12 (respectively) possessed identical composition, providing no evidence for selection of a chimeric CP with specific composition (data not shown). Remarkably, none of the recombinants had acquired the extreme N-terminal part of PVX-4 and only PVX-4-Rec3 had acquired the entire PVX-4 CP specific 11 aa insertion. The recombination sites detected in the recombinants were not distributed randomly over the CP gene and 3\textsuperscript{-ntr}, but clustered to five sites in proximity to the deletion in CP-UK3 (designated I–V in Supplementary Fig. S2), where 21 of 45 template switches were observed. The remaining sites appeared to be spread uniformly over the possible sequence range available for recombination.

**Precise homologous RNA recombination of PVX occurs preferentially at predicted stem–loop structures**

Assuming that recombination occurs during minus-strand synthesis, all parental sequences containing regions of possible RdRp template switch (Supplementary Fig. S2) were used for RNA secondary-structure prediction with the program Mfold 3.2 (Zuker, 2003). Analysis of regions I–V (Fig. 3) revealed a folding that leads to stem–loop structures. Remarkably, four of five recombination sites overlapped with the predicted loop and one template switch could be located within the loop. Analysis of the remaining 22 junction sites with more uniform distri- bution resulted in the prediction of recombination sites overlapping with a predicted loop structure in 17 cases (data not shown).

**DISCUSSION**

An experimental PVX recombination system developed previously (Draghici & Varrelmann, 2009) provides an important detection system for the identification of homo- logous recombinants with high probability and frequency within the functional PVX genome. It has been demonstrated here that this recombination system presents a powerful tool to study factors influencing the parameters of virus RNA recombination. PVX is a member of supergroup III (alpha-like virus) with rod-shaped particle morphology, and possesses a replicase protein differing in structure and functional domains from those encoded by the recombi- nation model plant viruses described so far (Goldbach & de Haan, 1994; Habili & Symons, 1989; Koonin & Dolja, 1993; reviewed by van der Heijden & Bol, 2002).

According to the current classification by Nagy & Simon (1997), replicative RNA recombination of model viruses characterized to date can be categorized into the following systems: similarity-essential (class 1, with either precise or aberrant end products), similarity non-essential (class 2) and similarity-assisted (class 3). In class 1, substantial sequence similarity is essential for recombination; RNA
secondary stem–loop structures are required for class 2 recombination to occur; class 3 recombination combines features of classes 1 and 2. Our results obtained and described above enable the assignment of the observed recombination parameters to an RNA recombination type defined by Nagy & Simon (1997).

Recombination rate and selection pressure
The application of intact PVX CP genes with variable sequence composition enabled determination of the number and positions of template switches leading to the formation of the intact CP gene by means of homologous recombination. As explained above, we have no evidence to specify whether the recombinant sequence was a product of virus plus- or minus-strand synthesis. When recombinants that represent a product of a single RdRp template switch were detected, these could be generated following minus-strand transcription initiation on the 3′-ntr of the intact transcript and completion of the viral genome synthesis after one template switch upstream of the CP deletion. Alternatively, recombination could have occurred during plus-strand synthesis following 5′-ntr replication initiation, as the virus deletion mutant itself was replication-competent (Draghici & Varrelmann, 2009). When an intact CP gene containing several mutations was offered for reconstitution of the defective viral gene, several recombinants were shown to be a product of at least two to seven template switches, indicating that, in this assay, PVX recombines at a rate higher than required to reconstitute the deletion. Most of the type X CP (PVX-14, -17, -18 and -847)-derived recombinants were unique, demonstrating that no strong selection pressure was exerted on the selection of a recombinant with a specific amino acid composition. This might be explained by the fact that these relatively few coding differences do not affect the functionality of the recombinant CP. To our knowledge, this high rate of template switching is a unique observation, compared with previous studies where only the number of template switches necessary for functional reconstitution of the gene

Fig. 3. RNA secondary structures of parental (+)-strand CP gene regions (PVX-4 or -UK3) where template switch occurred in proximity to the CP-UK3 deletion (I–V in Supplementary Fig. S2). The distance between mutations representing the possible region for RdRp template switch is indicated below each structure and highlighted in bold in the sequence. Nucleotide positions are according to the PVX-4 and -UK3 CP ORFs, respectively.
was reported (Allison et al., 1996; Borja et al., 1999; Gal-On et al., 1998; Varrelmann et al., 2000).

As the assay selects for precise homologous recombinants, all other recombination events, including those that lead to non-viable virus or debilitated mutants, remained undetected. It must be stressed that the recombination assay with high selection pressure does not allow deduction of the ability of PVX to recombine under natural conditions. Moreover, the viral genes and genomes were co-expressed in numerous cells via agroinfiltration. These artificial conditions might explain the frequent observation of recombinants, although evidence for this is currently lacking. The ability of the virus to recombine can only be depicted when minimal selection, i.e. in mixed infection, is applied (de Wispelaere et al., 2005), but recombinants do not occur frequently enough to study recombination parameters efficiently. However, in addition to the possible analysis of the recombination mechanism and parameters, the system might be useful to test virus-derived sequences for their ability to support recombination in virus-resistant transgenic plants. This will possibly help to evaluate biosafety issues before starting tedious plant-transformation experiments.

Recombination-mediated frequency of CP gene reconstitution is dependent on the length and similarity of homologous sequence overlap

The observation that the frequency of CP gene reconstitution was proportional to the length of the homologous sequence available for template switching is consistent with the findings of Nagy & Bujarski (1995). The relative recombination frequency observed in our assay obviously decreased linearly when the homologous sequence overlap was reduced from 600 to 50 nt. The conclusion may be drawn that the hypothetical heteroduplex formation (hybridization of parental strands), bringing homologous regions into proximity (Romanova et al., 1986), is dependent on the length of the homologous sequence, or is simply formed more frequently if the overlapping homologous region is longer. This confirmed the hypothesis that similarity is essential for precise reconstitution of the functional CP gene. This is strengthened by the observation that the sequence-divergent CP-4, although similar in length to CP-UK3, strongly reduced the frequency of CP gene reconstitution. More evidence is supplied by the observation that no recombination was detected on the scale of this assay when the overlap was reduced from 50 to 20 nt or below. On the other hand, the distance between mutations in CP-4–UK3 recombinants (Supplementary Fig. S2) varied between 2 and 21 nt, indicating that mismatches were tolerated at least to some extent in the homologous sequence stretch. Nagy & Bujarski (1995) provided evidence that nucleotide identity as short as 15 nt between parental BMV sequences supported efficient homologous recombination. In order to discern limits more precisely, deletion mutants with a more detailed resolution and higher number of independent repetitions are required. Taken together, these results prompted us to conclude that PVX homologous recombination does not occur frequently by a mechanism based on breakage and ligation, as described previously (see Introduction). To study replication-independent breakage and ligation recombination by host cell-encoded RNA ligase, two PVX deletion constructs would be required, separating the RdRp gene into two non-overlapping fragments. Although a critical length of homologous sequence overlap was determined with the system applied, non-replicative recombination by breakage and ligation cannot be excluded entirely.

Junction sites of homologous recombination are influenced by RNA secondary structure: evidence for similarity-assisted recombination

Intermolecular RNA secondary structures such as stem-loops have been shown to be important in directing non-homologous or similarity non-essential recombination sites in TCV, TBSV and cucumoviruses (Carpenter et al., 1995; Cascone et al., 1993; Chare & Holmes, 2006; de Wispelaere et al., 2005; Nagy & Simon, 1998; Nagy et al., 1998; Suzuki et al., 2003; White & Morris, 1994), whereas Nagy & Bujarski (1995) did not find a clear correlation between RNA secondary structure and homologous recombination sites. However, a high proportion (21 of 27) of independent recombination-junction sites between mutations were found in recombinants formed between X and Bi type CP genes and co-localized to predicted stem–loop structures. This finding provides strong evidence for an involvement of RNA secondary structure in precise homologous PVX recombination. These double-stranded structures might lead to RdRp stalling or template dissociation (Jager & Pata, 1999; Nagy & Simon, 1997) and possibly direct the template switch to the non-structured loop. This finding, however, is in contrast to the homologous recombination results obtained with BMV (Olthoorn et al., 2002), where a reduction in recombinant accumulation and no evidence for enhanced RdRp pausing and detachment were detected when stem–loop structures were introduced into the parental sequences. On the other hand, it has been demonstrated for BMV that intercistronic subgenomic CP promoter (sgPr) sequences, consisting of an AU-rich region and an RdRp-binding stem–loop structure, increase homologous recombination activity (Bruyere et al., 2000; Nagy et al., 1999; Wierzchoslawski et al., 2004). However, as the PVX CP ORF represents the most 3’ ORF in the genome, there is no evidence for the presence of an internal sgPr in the CP ORF. Considering that the experimental system selected amplification of only movement-competent CP genes and that only products of precise recombination events (no mismatch mutations in proximity to the junction sites) were detected, the results obtained allow classification of the precise recombination into the similarity-assisted group as defined by Nagy & Simon (1997). Another example for similarity-assisted recombination has been given (Cascone et al., 1993; Nagy
et al., 1998) for TCV satellite RNA; however, recombination hot spots were observed to cluster at the base of a hairpin motif. By using this recombination system, it should be possible to reduce successively the sequence similarity of intact genes offered for recombination, in order to analyse the similarity-dependent limits of precise homologous recombination. Thereby, the functional interchangeability of CP genes from more distant potexviruses could be investigated and the occurrence of aberrant homologous or even non-homologous recombination (if possible at all in PVX) could be forced, as described for other viruses (Carpenter & Simon, 1996; Figlerowicz, 2000; Figlerowicz et al., 1998; Hajjou et al., 1996). In addition, this experimental system can be adapted for the easy and straightforward characterization of recombination parameters of other viruses, where full-length cDNA clones for agroinfection or particle bombardment are available. The detection of recombinants between phylogenetic distant X and Bi type CP genes demonstrates the suitability of RNA detection of recombinants between phylogenetically distant X virus CP genes.

**Sequence analysis.** Sequence alignment was carried out by using the neighbour-joining algorithm in the CLUSTAL_X version 1.83 program (Thompson et al., 1997). The RNA secondary structures of parental (+) strand CP gene +3′-ntr sequences were predicted for PVX-UK3 and PVX-3 at 37 °C with 150 bp maximum distance between paired bases by using Mfold version 3.2 (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi) based on free-energy minimization (Mathews et al., 1999; Zuker, 2003).

**Clones and infectious transcripts.** The cauliflower mosaic virus (CaMV) 35S promoter-driven PVX plant expression plasmid cloned in a binary vector for agroinfection (pGR106; GenBank accession no. AY297843) (Jones et al., 1999; Lu et al., 2003) used in this study was supplied by David Baulcombe (Sainsbury Laboratory, Norwich, UK). The construction of a PVX (isolate UK3) plant expression vector expressing GFP, named 35S-PVX-GFP, was described previously (Draghici & Varrelmann, 2009).

**35S-PVX-GFP-derived mutants.** The identity of clones was confirmed via DNA sequencing by a commercial company (Eurofins MWG Operon, Martinsried, Germany). All recombinant plasmid manipulations were carried out as described by Sambrook et al. (1989) unless otherwise specified. The construction of a movement-defective PVX CP mutant with a 514–646 nt PvuI deletion, named 35S-PVX-GFP-ΔCP, was described previously (Draghici & Varrelmann, 2009). To generate CP deletion mutants with different lengths in 35S-PVX-GFP, the following strategy was applied. The 3′ end of the GFP gene, the sgPr and the 5′ end of the CP gene with variable lengths were PCR-amplified with specific primers using 35S-PVX-GFP as template. The upper primer CP-del-up (5′-AAATCTGCCCTTCTGAAGATGAT-3′) (in all primers, restriction sites are underlined and virus-specific sequences are in italics), homologous to a part of the GFP ORF, contained a Bsp119I restriction site for subsequent exchange in 35S-PVX-GFP and was used for the generation of all CP deletion mutants. Different lower CP primers (supplied with an Xhol restriction site) were used. Together with primer CP-del-up, this approach produced CP gene fragments of different lengths. The different PCR fragments were Bsp119I/Xhol-cloned into 35S-PVX-GFP, replacing most of the CP ORF and resulting in the generation of CP mutants with a 3′ end of fixed length (nt 653–711) and a 5′ end of variable length (0, nt 0–10, –20, –50, –101, –200, –300, –401, –500 and –600, respectively). The resulting clones were named 35S-PVX-GFP-CP-0, -10, -20, -50, -101, -200, -300, -401, -500 and -600, respectively (Fig. 1) and were subsequently applied in the recombination assay.

**Construction of binary vectors expressing complete CP gene+3′-ntr in vivo transcripts derived from different PVX isolates.** The exact CP ORF and complete 3′-ntr regions of PVX-UK3, PVX-14, PVX-17, PVX-847 and PVX-847 were RT-PCR-amplified from total RNA extracts (RNaseq; Qiagen) of virus-infected *N. benthamiana* plants. A PVX subtype Bi isolate, PVX-4 (group 4) (Santa Cruz & Baulcombe, 1995), was obtained and propagated in *N. benthamiana* leaf tissue, introducing flanking Sd1 and BamHI restriction sites. The following primers were used: CP-up (5′-TTAGAGGCTACATGTAGCAGCACAGTAGCAG-3′) and poly-T-down (Draghici & Varrelmann, 2009). In order to amplify the respective sequence from a PVX Bi type isolate, PVX-4, a derivative upper primer, CP4-up (5′-TTAGAGGCTACATGCTAGCAGCACAGTAGCAG-3′), was applied. The PCR fragments were Sall/BamHI-cloned into a binary vector, pBIn615s, under control of the enhanced CaMV 35S promoter and poly(A) terminator (Silhavy et al., 2002). Plasmid sequences were sequenced and named 35S-CP-UK3, -PVX-14, -PVX-17, -PVX-18, -PVX-847 and -PVX-4.

**Cloning and sequencing of CP gene and 3′-ntr from recombinant PVX isolates.** Upper leaves of *N. benthamiana* plants were selected for the recombination assay. Leaves displaying systemic infection, as monitored by GFP fluorescence, were used for total RNA extraction and RT-PCR amplification of the PVX CP gene and 3′-ntr, as described previously (Draghici & Varrelmann, 2009).

**Recombination assay.** In general, the recombination assay was performed as described previously (Draghici & Varrelmann, 2009). To produce PVX-GFP-UK3 with a recombinant CP gene, 35S-PVX-GFP-ΔCP was co-infiltrated with 35S-CP-UK3, -PVX-14, -PVX-17, -PVX-18, -PVX-847 or -PVX-4. Agroinfiltrated leaf patches were examined for GFP fluorescence with a hand-held long-wave UV lamp (Black Ray model B100 AP, 100 W, UV products) to visualize occurrence and movement of recombinant PVX–GFP. Photographs were taken using a digital camera (Nikon D50).

**ACKNOWLEDGEMENTS**

We kindly acknowledge Professor David Baulcombe for the gift of the pGR106 vector. D. Silhavy (Agricultural Biotechnology Center, C.

**METHODS**

**PVX strains.** To obtain CP genes with variable sequence similarity in comparison to the PVX-UK3 full-length clone (type X, group 3; Kagiwada et al., 2002; Kirkegaard & Baltimore, 1986), PVX type X isolates (Santa Cruz & Baulcombe, 1995) from different geographical origins (isolates PVX-14, PVX-17 (both Germany), PVX-18 (Afghanistan) and PVX-847 (Iran), kindly supplied by Stefan Winter (Plant Virus Department, German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany]) were obtained and propagated on *N. benthamiana* plants. A PVX subtype Bi isolate, PVX-4 (group 4) (Santa Cruz & Baulcombe, 1995), displaying higher sequence divergence from UK3, was kindly provided by Rene van der Vlugt (Plant Research International, Wageningen, the Netherlands).

**Sequence analysis.** Sequence alignment was carried out by using the neighbour-joining algorithm in the CLUSTAL_X version 1.83 program (Thompson et al., 1997). The RNA secondary structures of parental (+) strand CP gene +3′-ntr sequences were predicted for PVX-UK3 and PVX-3 at 37 °C with 150 bp maximum distance between paired bases by using Mfold version 3.2 (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi) based on free-energy minimization (Mathews et al., 1999; Zuker, 2003).
Godöllő (Hungary) is acknowledged for providing the pBIN61S plasmid. We thank M. Stevens for critically reading the manuscript. This research was financed by a fellowship (grant no. VA202/3-2) of the Deutsche Forschungsgemeinschaft.

REFERENCES


http://vir.sgmjournals.org


