Experimental and bioinformatic evidence that raspberry leaf blotch emaravirus P4 is a movement protein of the 30K superfamily

Chulang Yu,1† David G. Karlin,2,3† Yuwen Lu,1 Kathryn Wright,4 Jianping Chen1 and Stuart MacFarlane4

1Institute of Virology and Biotechnology, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, PR China
2Department of Zoology, Oxford University, South Parks Road, Oxford OX1 3PS, UK
3Division of Structural Biology, Henry Wellcome Building, Roosevelt Drive, Oxford OX3 7BN, UK
4James Hutton Institute, Cell and Molecular Sciences Group, Invergowrie, Dundee DD2 5DA, UK

Emaravirus is a recently described genus of negative-strand RNA plant viruses. Emaravirus P4 protein localizes to plasmodesmata, suggesting that it could be a viral movement protein (MP). In the current study, we showed that the P4 protein of raspberry leaf blotch emaravirus (RLBV) rescued the cell-to-cell movement of a defective potato virus X (PVX) that had a deletion mutation in the triple gene block 1 movement-associated protein. This demonstrated that RLBV P4 is a functional MP. Sequence analyses revealed that P4 is a distant member of the 30K superfamily of MPs. All MPs of this family contain two highly conserved regions predicted to form β-strands, namely β1 and β2. We explored by alanine mutagenesis the role of two residues of P4 (Ile106 and Asp127) located in each of these strands. We also made the equivalent substitutions in the 29K MP of tobacco rattle virus, another member of the 30K superfamily. All substitutions abolished the ability to complement PVX movement, except for the I106A substitution in the β1 region of P4. This region has been shown to mediate membrane association of 30K MPs; our results show that it is possible to make non-conservative substitutions of a well-conserved aliphatic residue within β1 without preventing the membrane association or movement function of P4.

INTRODUCTION

The genus Emaravirus is a very recently established grouping of negative-strand RNA plant viruses for which European mountain ash ringspot-associated virus (EMARAV) is the type species (Mielke & Muehlbach, 2007). Other accepted species in the genus are Fig mosaic virus and Rose rosette virus, and putative members are maize red stripe virus (MRSV; also referred to as High Plains virus), pigeon pea sterility mosaic virus (PPSMV), raspberry leaf blotch virus (RLBV) and redbud yellow ringspot virus (RYRSV). Some of these viruses have been shown to have a roughly spherical morphology of 80–100 nm diameter with a surrounding envelope contained within vesicles thought to be derived from host membranes (Ebrahim-Nesbat & Izadpanah, 1992). Emaraviruses are reported to be transmitted by eriophyid mites of various species, and EMARAV has been detected within the body cavity of the mite, suggesting that it may circulate within the mite and, similar to the enveloped plant tospoviruses, might also multiply within the mite (Mielke-Ehret et al., 2010). Purification of emaraviruses from infected plants is complicated by the low titre and enveloped nature of the virus particles, and this has made it difficult to determine the definitive genetic composition of these viruses. Thus, EMARAV and rose rosette virus (RRV) are reported to have four genomic RNA segments (Mielke & Muehlbach, 2007; Laney et al., 2011). Fig mosaic virus (FMV) was initially reported also to have four RNAs, which have now been supplemented by another two (Elbeaino et al., 2009; Ishikawa et al., 2012). RLBV was reported initially to have five RNA segments but is now known to have at least eight (McGavin et al., 2012; S. MacFarlane, unpublished data).

The emaravirus RNA segments each encode a single putative protein, three of which are predicted to function as an RNA-dependent RNA polymerase (encoded by RNA1), a glycoprotein precursor protein (RNA2) and a nucleocapsid protein (RNA3) (Mielke-Ehret & Muehlbach, 2010).
2012). The functions of the other putative viral proteins cannot be predicted confidently from inspection of their amino acid sequences; however, at least one of them would be expected to be a movement protein (MP), responsible for spread of the virus between cells. Plant viruses move from cell to cell via plasmodesmata (PD), which are complex, membrane-lined channels that penetrate the cell wall. Many virus-encoded MPs associate with the plant’s endoplasmic reticulum (ER) and plasma membranes, thereafter locating to and modifying PD to increase the size limit for passage of molecules through the PD (Schoelz et al., 2011). Plant virus MPs belong to several different structural classes, and one of the largest groups is the 30K MP superfamily whose members are structurally related to the tobacco mosaic virus (TMV) 30K MP (Koonin et al., 1991; Mushgein & Koonin, 1993; Melcher, 2000). Experiments have shown that 30K family MPs can be interchangeable, so that, for example, the cell-to-cell and systemic movement of alfalfa mosaic virus can be brought about by the MPs of viruses from five other genera (Sánchez-Navarro et al., 2006; Fajardo et al., 2013).

The cell-to-cell movement of potato virus X (PVX) and some other viruses utilizes three viral proteins, comprising the triple gene block (TGB), which have no structural similarity to 30K family proteins (Verchot-Lubicz et al., 2010). However, underlying shared aspects of the movement process have been revealed by the finding that movement of TGB1-deficient PVX can be rescued by the 30K MP (Morozov et al., 1997) and that TGB1 can rescue movement of 30K MP-deficient Odontoglossum ringspot virus, a tobamovirus related to TMV (Ajjikuttira et al., 2005).

In a previous study, we showed that the RLBV 42 kDa P4 protein fused to GFP or monomeric red fluorescent protein (mRFP) localized to the plasma membrane and co-localized with TMV 30K MP to PD in the cell wall (McGavin et al., 2012). These results suggested that P4 could be a virus MP and prompted us to investigate this possibility using sequence analysis, site-directed mutagenesis and movement complementation approaches.

RESULTS AND DISCUSSION

RLBV P4 complements movement-defective PVX

No infectious clone system is available for the genetic analysis of any emaravirus, including RLBV, so we could not directly investigate the possible role of P4 in the movement of RLBV. As an alternative approach, we decided to examine whether the RLBV P4 protein could rescue the cell-to-cell movement of PVX expressing GFP but with an in-frame deletion within the TGB1 gene that prevents PVX cell-to-cell movement (Bayne et al., 2005).

To confirm that the PVX movement complementation system could work in our hands, we used as a known MP the 29K protein from tobacco rattle virus (TRV), a close homologue of the TMV 30K MP. At 3 days post-infiltration (p.i.) of leaves with the movement-deficient PVX (PVX-GFP-∆MP), GFP fluorescence was detected only in isolated epidermal cells, showing that the virus had replicated in the initial infected cell but had not moved to adjacent cells (Fig. 1, left panel). No movement was detected even at later times (more than 1 week). In contrast, when PVX-GFP-∆MP was co-infiltrated with a TRV 29K protein construct, by 4 days p.i. expanding foci of GFP fluorescence, where the mutant virus had moved from the initial infected cell into surrounding cells, could be seen both by confocal microscopy and by eye when the leaves were illuminated using a UV lamp (Fig. 1, centre panel). These results showed that, as expected, the TRV 29K protein could complement the movement deficiency of the PVX-GFP-∆MP TGB1 mutant.

Having established that our experimental set-up worked, we conducted similar experiments on the RLBV P4 protein. P4 was also able to complement movement of PVX-GFP-∆MP, demonstrating for the first time its ability to function as a virus MP (Fig. 1, right panel). An RLBV P4-mRFP fusion protein also complemented movement of PVX-GFP-∆MP (data not shown), demonstrating that making fusions to the C terminus of P4 did not interfere with its movement function.

Emaravirus P4 MPs belong to the 30K superfamily

A multiple sequence alignment of the P4 proteins from four emaraviruses is presented in Fig. 2. The protein described as P4 in EMARAV (GenBank accession no. YP_003104766) has a different length to other P4 proteins, a different secondary structure and no detectable sequence similarity (see below) and was therefore not included in the alignment. In all other emaraviruses, P4 comprised a predicted signal peptide (aa 2–21 in RLBV), followed by an N-terminal moiety composed mainly of predicted β-strands with some interspersed α-helices (approx. aa 37–238), and by a C-terminal moiety composed of long α-helices predicted to form one or several coiled coils (aa 244–348) (Fig. 2).

We noted that the secondary structure pattern of the central region of P4 (aa 82–210, boxed in Fig. 2) was similar to the consensus secondary structure of the 30K superfamily of plant virus MPs (Melcher, 2000). The 30K MPs have little overall sequence similarity, but their central domains all have the same predicted secondary structure pattern (Fig. 3), formed by a series of α-helices and β-strands, named (from N to C terminus) αA, β1, β2, αB and β3–β7. Some 30K MPs have additional predicted α-helices, such as αC between strands β5 and β6, and/ or αD downstream of strand β7 ((Melcher, 2000; and our observations). There are minor differences between the secondary structure of the central part of P4 and the consensus secondary structure of 30K MPs: in P4, the beginning of the region corresponding to predicted strand β1 in 30K MPs was predicted to form a short α-helix, αA’
In addition to having a conserved secondary structure, 30K MPs contain a short region conserved in sequence, corresponding to strands $\beta_1$ and $\beta_2$; as can be seen in Fig. 3, this region contains several conserved hydrophobic or aliphatic positions, and a conserved D/N at the C terminus of $\beta_2$ (substituted by Y in nepoviruses and nucleorhabdoviruses; Koonin et al., 1991; Mushegian & Koonin, 1993; Melcher, 2000). In addition, there is often a small residue (such as G) upstream of $\beta_6$. The corresponding residues were also conserved in P4 (Fig. 2). In particular, the conserved D/N at the C terminus of $\beta_2$ corresponded to D127 of RLBV P4 (Figs 2 and 3), and there was a conserved small residue (G or A) upstream of $\beta_6$ in P4 (Fig. 2).

Neither the program BLAST nor the more sophisticated PSI-BLAST (Altschul et al., 1990, 1997) could detect significant sequence similarity between emaravirus P4 and 30K MPs or other viral or cellular proteins (Laney et al., 2011; Ishikawa et al., 2013). However, more powerful similarity detection methods have been developed recently, which rely on profile–profile comparisons, such as HHpred (Hildebrand et al., 2009), HHblits (Remmert et al., 2012) HHalign (Biegert et al., 2006), FFAS (Jaroszewski et al., 2011) and WebPRC (Brandt & Heringa, 2009). In brief, a sequence profile is a representation of a multiple sequence alignment that contains information about which amino acids are ‘tolerated’ at each position of the alignment, and with which probability. Comparing two profiles is a much more sensitive method than comparing two sequences, because the profiles contain information about how the sequences can evolve, and can thus identify weak similarities that remain after the two sequences have evolved apart (Dunbrack, 2006; Söding & Remmert, 2011; Karlin & Belshaw, 2012).

We compared the multiple sequence alignment of emaravirus P4 proteins with an alignment of representative MPs of the 30K superfamily using HHalign (Biegert et al., 2006) (see Methods). HHAlign reported that the region of P4 corresponding to strands $\beta_1$–$\beta_6$ had statistically significant similarity ($E = 1.7 \times 10^{-5}$) with the $\beta_1$–$\beta_6$ region of 30K MPs, indicating that they are homologous. Only the region $\beta_1$–$\beta_2$ of P4 had good sequence conservation with other 30K MPs (Fig. 2); the reason why HHalign reported a longer region of similarity, extending to $\beta_6$, was thus probably that HHalign detects similarity in secondary structure and not only in sequence (Biegert et al., 2006).

In conclusion, the combination of similarity in secondary structure between P4 and 30K MPs, of their similar amino acid motifs and of their statistically significant sequence similarity reliably indicate that P4 is a member of the 30K superfamily.

**Particular features of emaravirus P4 proteins**

The fact that simple sequence searches using, for instance, PSI-BLAST, detected no similarity between P4 and any other 30K MP indicated that P4 is only distantly related to other 30K MPs. Its most striking feature was the predicted coiled coil downstream of the central conserved domain (Fig. 2). This feature is reminiscent of the MP of cauliflower mosaic virus (CaMV), which trimerizes through a C-terminal coiled-coil region. Interestingly, this region is thought to form a heteromeric coiled coil with the CaMV-encoded virion-associated protein (Stavolone et al., 2005). Our results indicated that the C terminus of P4 is probably accessible on the surface of the protein, as large C-terminal fusion tags such as GFP did not destroy the movement function of P4. Further experiments should thus investigate whether the coiled-coil region of P4 promotes binding to other viral or cellular proteins.

We noted that the protein currently annotated as ‘P4’ in EMARAV is completely unrelated in sequence to the other emaravirus P4 proteins or to any known emaravirus protein. Therefore, we suggest that the RNA encoding the MP of EMARAV remains to be identified. Indeed, there probably remain several genomic RNAs to be discovered in most emaraviruses (see Introduction).
Mutation of residues conserved in the 30K superfamily affect the movement function of P4 and its localization to PD

To characterize further the RLBV P4 protein, we mutated two of the residues that are conserved throughout the 30K superfamily, one within strand $\beta_1$ and the other within strand $\beta_2$ (in bold above the alignment in Fig. 2), and examined the effects of these mutations on the movement function of the protein. At the same time, we mutated the equivalent residues in the TRV 29K MP. Three mutants of each protein were created: for RLBV P4 (Table 2), substitutions I106A (replacement of Ile in $\beta_1$ with Ala), D127A (replacement of Asp in $\beta_2$ with Ala) and I106A D127A (both mutations combined), and for TRV 29K, substitutions V81A (replacement of Val in $\beta_1$ with Ala), D103A (replacement of Asp in $\beta_2$ with Ala) and V81A D103A (both mutations combined). GFP and mRFP fusions of each of the mutant P4 and 29K proteins produced observable fluorescence when the constructs were agroinfiltrated into Nicotiana benthamiana leaves, showing that all of the mutant proteins could be expressed and accumulate in plant cells, although the TRV mutant proteins produced weaker fluorescence than the WT 29K protein (data not shown).

Only the P4 $\beta_1$ I106A mutant was able to rescue cell-to-cell spread of the movement-deficient PVX-GFP-D MP (Fig. 4, diagram).

Fig. 2. Alignment of emaravirus P4 proteins. Consensus secondary structure elements predicted by the software Promals (Pei et al., 2007) are shown below the alignment. The central part of P4 that is similar to 30K MPs is boxed and its secondary structure elements are named according to the nomenclature of 30K MPs (Melcher, 2000). GenBank accession numbers and full virus names are listed in Table 1. See Fig. 3 for definitions of the colour coding.
Fig. 3. Alignment of the β1−β2 region conserved in sequence in the MPs of all members of the 30K superfamily. Conventions are as in Fig. 2. The text to the left indicates virus names and genus (in capitals). The position of strands β1 and β2 corresponds to consensus predictions; because the protein sequences aligned are very distant, this position should be considered an ‘average’, which may vary among members of the 30K superfamily. Residues highlighted in bold indicate where mutations have been introduced within this region (see text and Table 2). The location of the residues of RLBV P4 chosen for mutation within β1 and β2 is shown above the alignment.
These results demonstrated that the conserved residue D127, located in strand $b_2$, is critical for both the movement activity and localization to PD of P4, whereas, in contrast, the conserved aliphatic residue at position 106 of strand $b_1$ in P4 is not indispensable to either of these functions.

### Comparison of effects of mutations in $b_1$ and $b_2$ in MPs of the 30K superfamily

In all MPs of the 30K superfamily in which the conserved D residue of strand $b_2$ has been experimentally substituted, the substitution resulted in loss of virus movement (Table 2). This includes the MPs of RSV, cowpea mosaic virus (CPMV), TSWV and CaMV (Thomas & Maule, 1995; Bertens et al., 2000; Li et al., 2009; Zhang et al., 2012) (in MPs that form tubules at the cell wall, such as CPMV and TSWV, the substitution also abolished the formation of tubules). Mutation of the conserved D residue of emaravirus RLBV P4 also abolished movement, and thus this D residue could be considered a hallmark to experimentally characterize the 30K superfamily. The loss of movement was unlikely to be due to a major effect on the structural integrity of P4, as the protein was expressed in normal quantities and still localized at the membranes; instead, it perhaps resulted in a loss of interaction with a PD protein(s) that remains to be identified.

In contrast to the conserved D residue of strand $b_2$, the impact of substitutions of the well-conserved aliphatic position in strand $b_1$ varies (Table 2). For the RLBV P4 protein, the Ile106 residue was not critical for movement capability or PD localization. In the same manner, a (conservative) substitution of the equivalent Leu with Val or Ile resulted in improved TMV movement (Toth et al., 2002; Kawakami et al., 2003). In contrast, substitutions of the equivalent $b_1$ residue abolished virus movement in TRV (this study), and in CPMV, Chinese wheat mosaic virus (CWMV) and Prunus necrotic ringspot virus (PNRSV), as well as tubule formation in CPMV (Bertens et al., 2000, Martinez-Gil et al., 2009; Andika et al., 2013), whilst deletion of part of the $b_1$ region prevented TMV movement (Fujiki et al., 2006).

Interestingly, $b_1$ is part of a region experimentally shown to associate with membranes, although the precise mechanism of membrane association of 30K MPs is still a matter of debate. On the one hand, a series of experiments on TMV 30K concluded that it was most probably an integral membrane protein with two $a$-helical transmembrane segments (although a tight association with membranes without transmembrane domains could not be excluded; Brill et al., 2000; Fujiki et al., 2006). The two transmembrane segments correspond to the regions predicted as strands $b_1$ and $b_6$ in the study of Melcher (2000), which is the nomenclature followed in the present study. In contrast, another detailed study of the MP of an ilavirus, PNRSV, found that it had no transmembrane segment but was instead tightly associated with the membrane through an $a$-helical region encompassing the predicted strand $b_1$. This discrepancy is unexpected; rather, given the high sequence conservation within region $b_1$ (Fig. 3), one would expect the mechanism of its association with membranes to be similar in all 30K MPs.

<table>
<thead>
<tr>
<th>Family</th>
<th>Genus</th>
<th>Virus species</th>
<th>Abbreviation</th>
<th>Protein name</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betaflexiviridae</td>
<td>Capillovirus</td>
<td>Apple stem grooving virus</td>
<td>ASGV</td>
<td>36K</td>
<td>CBY84302.1</td>
</tr>
<tr>
<td></td>
<td>Trichovirus</td>
<td>Apple chlorotic leaf spot virus</td>
<td>ACLSV</td>
<td>50K, MP</td>
<td>NP_040552.1</td>
</tr>
<tr>
<td>Bromoviridae</td>
<td>Alfamovirus</td>
<td>Alfalfa mosaic virus</td>
<td>AIMV</td>
<td>MP</td>
<td>BAE94265.1</td>
</tr>
<tr>
<td></td>
<td>Brome mosaic virus</td>
<td></td>
<td>BMV</td>
<td>3a</td>
<td>P03603.1</td>
</tr>
<tr>
<td></td>
<td>Cucumovirus</td>
<td>Cucumber mosaic virus</td>
<td>CMV</td>
<td>3a</td>
<td>AEK69523.1</td>
</tr>
<tr>
<td></td>
<td>Ilarvirus</td>
<td>Prunus necrotic ringspot virus</td>
<td>PNRSV</td>
<td>MP</td>
<td>NP_733825.1</td>
</tr>
<tr>
<td>Bunyaviridae</td>
<td>Tospovirus</td>
<td>Tomato spotted wilt virus</td>
<td>TSWV</td>
<td>NSm</td>
<td>BADS1469.1</td>
</tr>
<tr>
<td>Caulimoviridae</td>
<td>Caulimovirus</td>
<td>Cauliflower mosaic virus</td>
<td>CVMV</td>
<td>MP</td>
<td>AAA62371.1</td>
</tr>
<tr>
<td>Geminiviridae</td>
<td>Begomovirus</td>
<td>Bean golden mosaic virus</td>
<td>BGMV</td>
<td>BL1</td>
<td>AAA4617.1</td>
</tr>
<tr>
<td>Rhabdoviridae</td>
<td>Nucleorhabdovirus</td>
<td></td>
<td>RYSV</td>
<td>P3</td>
<td>Q98663.1</td>
</tr>
<tr>
<td>Secoviridae</td>
<td>Comovirus</td>
<td>Cowpea mosaic virus</td>
<td>CPMV</td>
<td>MP</td>
<td>NP_734000.1</td>
</tr>
<tr>
<td></td>
<td>Nepovirus</td>
<td>Grapevine fanleaf virus</td>
<td>GFLV</td>
<td>P2</td>
<td>ABB91827.1</td>
</tr>
<tr>
<td>Tombusviridae</td>
<td>Tombusvirus</td>
<td>Tomato bushy stunt virus</td>
<td>TBSV</td>
<td>p22</td>
<td>PS0631.1</td>
</tr>
<tr>
<td>Virgaviridae</td>
<td>Furovirus</td>
<td>Chinese wheat mosaic virus</td>
<td>CWMV</td>
<td>37K</td>
<td>NP_059515.1</td>
</tr>
<tr>
<td></td>
<td>Tobamovirus</td>
<td>Tobacco mosaic virus</td>
<td>TMV</td>
<td>30K</td>
<td>CA98976.1</td>
</tr>
<tr>
<td></td>
<td>Tobravirus</td>
<td>Tobacco rattle virus</td>
<td>TRV</td>
<td>29K</td>
<td>ACX54060.1</td>
</tr>
<tr>
<td>Unclassified; related to</td>
<td>Unbravirus</td>
<td>Tobacco bushy top virus</td>
<td>TBTB</td>
<td>ORF4</td>
<td>CAX62135.1</td>
</tr>
<tr>
<td></td>
<td>Emaravirus</td>
<td>European mountain ash ringspot-associated virus</td>
<td>EMARAV</td>
<td>P4</td>
<td>YP_003104766</td>
</tr>
<tr>
<td></td>
<td>Fg mosaic virus</td>
<td></td>
<td>FMV</td>
<td>P4</td>
<td>CAX32469.1</td>
</tr>
<tr>
<td></td>
<td>Raspberry leaf blotch virus</td>
<td>RLBV</td>
<td>P4</td>
<td>CBZ42027.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Redbud yellow ringspot virus</td>
<td>RYRSV</td>
<td>P4</td>
<td>AEO88257.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rose rosette virus</td>
<td>RRV</td>
<td>P4</td>
<td>YP_004327592.1</td>
<td></td>
</tr>
<tr>
<td>Tenuivirus</td>
<td>Rice stripe virus</td>
<td>RSV</td>
<td>pc4, NSvc4</td>
<td>ACSR7055.1</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Virus names and GenBank accession numbers of the proteins studied
Also, curiously, we noted that this region is consistently predicted as a β-strand in all 30K MPs rather than as an α-helix (Melcher, 2000; and our observations). Thus, on the one hand, the region we have referred to as β1 has been shown to associate with membranes in several studies of distantly related MPs; on the other hand, the precise mechanism of membrane association of 30K MPs (integral or peripheral) and the actual secondary structure adopted by this region await confirmation.

Our preliminary characterization of P4 shows that it is possible to make non-conservative substitutions of an aliphatic residue that is well conserved in β1 without disrupting either membrane association, plasmodesmal localization or preventing the movement function of P4; we hope further studies will uncover mutants that dissociate different functions of P4 and reveal the role(s) of this region.

Comparison with previous studies

Whilst this manuscript was in preparation, a study investigating the function of the P4 protein of another emaravirus, FMV, was published (Ishikawa et al., 2013). It showed that FMV P4 was able to complement the cell-to-cell spread of a movement-defective PVX and formed tubule-like structures at PD. Our study supports their findings but also provides bioinformatic evidence that P4 is related to a large superfamily of known MPs, and compares the effect of substitutions of conserved residues of P4 with studies carried out on other 30K family MPs.

Previous studies based on conserved domain search queries (Marchler-Bauer & Bryant, 2004; Marchler-Bauer et al., 2013) suggested that the C terminus of emaravirus P4 is similar to the DnaK peptide-binding domain (Laney et al., 2011; Ishikawa et al., 2013). The DnaK peptide-binding domain has a fold composed of a β-sheet subdomain followed by an α-helical subdomain (Zhu et al., 1996), and thus its secondary structure arrangement resembles that of P4 (Fig. 3). However, the C terminus of P4 is strongly predicted to form a coiled coil, which would be incompatible with the arrangement of the four α-helices of the last subdomain of DnaK (Zhu et al., 1996). We compared a multiple sequence alignment of the DnaK peptide-binding domain (corresponding to the last 219 aa of the PFAM family HSP70) with an alignment of emaravirus P4. HHalign reported a weak similarity (E=0.015) between a short region of P4 within its coiled-coil domain (aa 316–348 in RBLV P4) and the last two helices of the DnaK peptide-binding domain. Thus, the suggested similarity between P4 and DnaK is probably not authentic, as it is not statistically significant and occurs within the coiled-coil region, which is known to provoke spurious hits in similarity searches with unrelated helical proteins (Ferron et al., 2006; Gruber et al., 2006).

METHODS

Cloning. The RLBV P4 and TRV 29K genes were amplified by reverse transcriptase-PCR and sequenced using standard methods. Mutations were introduced into the genes using a QuickChange Site-Directed Mutagenesis kit (Stratagene), following the supplier’s instructions.
Movement complementation assay. A binary plasmid carrying the movement-deficient infectious clone of PVX-GFP-AMP was transformed into *Agrobacterium tumefaciens* strain GV3101. Overnight bacterial cell cultures were resuspended in infiltration buffer (Voinnet et al., 2003) to an OD$_{595}$ of 0.1 and then diluted 1:10,000 before infiltration into *N. benthamiana* plants.

The RLBV P4 and TRV 29K MP genes were cloned using Gateway technology (Invitrogen) into the binary plasmid pMDC32 (Curtis & Grossniklaus, 2003) and transformed into *A. tumefaciens* strain GV3101. Overnight cultures of these constructs were resuspended in infiltration buffer to an OD$_{595}$ of 0.3, combined with the diluted resuspension of the PVX mutant and infiltrated into *N. benthamiana* leaves.

Confocal microscopy. Detached leaves from plants that had previously been infiltrated with *Agrobacterium* cultures carrying the various virus and MP constructs were examined using a Leica TCS-SP2 AOBS confocal laser-scanning microscope (Leica Microsystems) fitted with a Leica HCX APO $\times 63/0.9$ W water-dipping lens. GFP was imaged singly or sequentially in combination with mRFP or mOrange at the following wavelengths: GFP (green): excitation 488 nm, emission 500–530 nm; mRFP or mOrange

---

**Fig. 5.** Effect of $\beta_1$ and $\beta_2$ mutations on the subcellular localization of RLBV P4–GFP. Images are leaves infiltrated with: WT RLBV P4 (top row), P4 $\beta_1$ mutant (second row), P4 $\beta_2$ mutant (third row) or P4 $\beta_1$ + $\beta_2$ mutant (bottom row). All leaves were co-infiltrated with mOrange-LT16b membrane marker. The P4–GFP signal is shown in the first column, the mOrange-LT16b signal in the second column and the merged GFP+mOrange signals in the third column. Arrows indicate P4–GFP located to PD. Bars, 50 μm (rows 1, 2 and 4); 25 μm (row 3).
Table 2. Mutation of β1 and β2 regions of diverse 30K family MPs from different viruses

<table>
<thead>
<tr>
<th>Genus</th>
<th>Virus species</th>
<th>Protein name*</th>
<th>Substitution(s)†</th>
<th>Region</th>
<th>Effect of the mutation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caulimovirus</td>
<td>Cauliflower mosaic virus</td>
<td>P1</td>
<td>Series of 3 aa deletions</td>
<td>β1 and β2</td>
<td>No plant infection</td>
<td>Thomas &amp; Maule (1995)</td>
</tr>
<tr>
<td>Comovirus</td>
<td>Cowpea mosaic virus</td>
<td>48K MP</td>
<td>IE (aa 239–240)→AA§</td>
<td>β1</td>
<td>No tubule formation; no movement; lowered membrane localization</td>
<td>Bertens et al. (2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VD (aa 259–260)→AA§</td>
<td>β2</td>
<td>No tubule formation; no movement; lowered membrane localization</td>
<td></td>
</tr>
<tr>
<td>Enaravirus</td>
<td>Raspberry leaf blotch virus</td>
<td>P4</td>
<td>I106A</td>
<td>β1</td>
<td>No effect on viral movement</td>
<td>Bertens et al. (2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D127A</td>
<td>β2</td>
<td>Loss of localization to PD and of movement</td>
<td>This study</td>
</tr>
<tr>
<td>Furovirus</td>
<td>Chinese wheat mosaic virus</td>
<td>37K</td>
<td>VLI (aa 84–86)→PPP</td>
<td>β1</td>
<td>Loss of localization to PD and of movement; altered accumulation in the ER</td>
<td>Andika et al. (2013)</td>
</tr>
<tr>
<td>Ilarvirus</td>
<td>Prunus necrotic ringspot virus</td>
<td>MP</td>
<td>P94A, Q97All</td>
<td>β1</td>
<td>Altered glycosylation pattern of MP, and P96A mutant prevented movement</td>
<td>Martinez-Gil et al. (2009)</td>
</tr>
<tr>
<td>Tenuivirus</td>
<td>Rice stripe virus</td>
<td>pc4</td>
<td>D135A</td>
<td>β2</td>
<td>Loss of movement</td>
<td>Zhang et al. (2012)</td>
</tr>
<tr>
<td>Tospovirus</td>
<td>Tomato spotted wilt virus</td>
<td>NSm</td>
<td>D154A</td>
<td>β2</td>
<td>Loss of tubule formation and of movement</td>
<td>Li et al. (2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No movement and reduced membrane association</td>
<td>Fujiki et al. (2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Retains PD localization, no ER association, no movement</td>
<td>Kahn et al. (1998)</td>
</tr>
<tr>
<td>Tobamovirus‡</td>
<td>Tobacco mosaic virus</td>
<td>30K</td>
<td>LAG (aa 69–71) deletion</td>
<td>β1</td>
<td>Improved virus movement</td>
<td>Toth et al. (2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VSV (aa 88–90) deletion</td>
<td>β2</td>
<td>No movement</td>
<td></td>
</tr>
<tr>
<td>Tobravirus</td>
<td>Tobacco rattle virus</td>
<td>29K</td>
<td>L72V, L72I</td>
<td>β1</td>
<td>Improved virus movement</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>V81A</td>
<td>β1</td>
<td>Loss of movement</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>D103A</td>
<td>β2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*GenBank accession numbers for the corresponding proteins are in Table 1.
†The substitutions are in bold in Fig. 2.
‡The majority of studies have been carried out on TMV 30K MP and only a small selection are detailed here.
§Numbering is according to the mature peptide MP described in the NCBI reference genome, and differs by 117 aa from the location stated in the publication of Bertens et al. (2000) where the substitutions are IE(aa 122–123)→AA and VD(aa 142–143)→AA.
||Numbering is according to the MP described in the NCBI reference genome and differs by 2 aa from the publication of Martinez-Gil et al. (2009) where the substitutios are P96A and Q99A.
(magenta): excitation 561 nm, emission 590–630 nm. Images are presented as either single sections or as maximum-intensity projections of multiple-layered stacks. Images were assembled and edited using Adobe Photoshop CS version 8.0.

**Complementation of PVX-GFP-ΔMP movement by P4 proteins.** Leaves infiltrated with PVX-GFP-AMP in combination with the WT or mutant P4–mRFP fusion proteins were examined over a period of 9 days p.i. by confocal microscopy. The treatment effect on the area of infection foci with WT or β1 mutant P4 proteins was further examined at 7 days p.i. Whole leaves, illuminated with light at 365 nm provided by a Blak-Ray hand-held lamp (UVP Products), were imaged using a Canon EOS 350D camera fitted with an EFS 60 mm lens. The area of individual infection sites was measured, after thresholding, using Image J software.

Differences in foci area between the two treatments were examined using ANOVA, with treatment and leaf as factors and leaf nested within treatment. In order to have a balanced design, the number of foci per leaf was standardized to 112 for each of the five leaves per treatment (using one leaf per plant). A natural log transformation was needed to normalize the foci area.

**Bioinformatic analyses.** The accession numbers of the sequences of emaravirus P4 and other MPs analysed in this study, as well as the abbreviations of species names, are listed in Table 1. Prior to homology searches, we carried out good-practice checks such as detecting the presence of coiled coils, low-complexity sequences, alignment-based searching of public domain databases. For more powerful significance. We used BLAST and PSI-BLAST (Altschul et al., 1990). Basic local alignment search tool. J Mol Biol 215, 403–410. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25, 3389–3402.


**ACKNOWLEDGEMENTS**

This work was funded by the Scottish Government Rural and Environment Science and Analytical Services (RESAS) and supported by the Wellcome Trust grant number 090005 to D.K. This work is part of a collaborative project on Plant Virology between the Zhejiang Academy of Agricultural Sciences (ZAAS) and James Hutton Institute, funded by ZAAS. C.Y. and J.C. are supported by the Program for Zhejiang Leading Team of Science and Technology Innovation (2011R09032-10), the Program for Leading Team of Agricultural Research and Innovation of Ministry of Agriculture, China, the Modern Agricultural Biotechnology and Crop Disease Control from Key Subject Construction Program of Zhejiang Province, China, and the Project of Molecular Mechanism of Plant Defence to Pests and Diseases (2012CB722504) from the Ministry of Science and Technology, China, the State Basic Research Program of China (2010CB126203, 2012CB722504), the International Science & Technology Cooperation Program of China (2012DFA30900) and the China Agriculture Research System (CARS-3-1) from the Ministry of Agriculture of the People’s Republic of China. We thank P. Boevink for providing the mOrange-LTI6b construct and P. J. Wright for advice on experimental design.

**REFERENCES**


