



HAL
open science

Virulence domain of the RYMV genome-linked viral protein VPg towards rice rymv1-2-mediated resistance

Eugénie Hébrard, Agnès Pinel-Galzi, Denis Fargette

► **To cite this version:**

Eugénie Hébrard, Agnès Pinel-Galzi, Denis Fargette. Virulence domain of the RYMV genome-linked viral protein VPg towards rice rymv1-2-mediated resistance. *Archives of Virology*, 2008, 19, pp.2913 - 1164. 10.1007/s00705-008-0087-9 . hal-01522140

HAL Id: hal-01522140

<https://hal.science/hal-01522140>

Submitted on 12 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Virulence domain of the RYMV Genome-Linked Viral Protein VPg towards rice *rymv1-2*-mediated resistance

Eugénie Hébrard, Agnès Pinel-Galzi, Denis Fargette

UMR RPB 186, Institut de Recherche pour le Développement (IRD), BP 64501, 34394 Montpellier cedex 5, France.

Author for correspondence: E. Hébrard; Telephone: +33 (0)4 67 41 62 89; Fax: +33 (0)4 67 41 62 83; E-mail: hebrard@mpl.ird.fr

SUMMARY

Virulent variants of *Rice yellow mottle virus* (genus *Sobemovirus*) can emerge on the highly resistant rice cultivars Gigante and Bekarosaka. Non-synonymous mutations responsible of the breakdown of the recessive resistant gene *rymv1-2* were located in the VPg after determination of its termini in the polyprotein P2a. The secondary structure of this protein was predicted to include a central α -helix. The two major amino acids related to virulence are located in the same side of this helix. The 3D-topology and the biochemical properties of virulence mutations both suggested a direct site-to-site interaction between RYMV VPg and rice eIF(iso)4G encoded by *rymv1*.

Rice yellow mottle virus (RYMV) of the genus *Sobemovirus* [7] is present in all rice-growing African countries where it causes high yield losses [9]. Two *Oryza sativa indica* cultivars Gigante and Bekarosaka and a few *O. glaberrima* cultivars Togs expressed high resistance towards RYMV [1]. However, several RYMV isolates of different strains were able to overcome the high resistance [13]. Mutations associated with virulence towards cvs. Gigante and Bekarosaka were restricted to five sites in the RYMV genome [13]. They are located in a stretch of 15 amino acids in the polyprotein P2a which contains sequence motifs of the serine protease and of the VPg (Viral Protein genome-linked). The sites and the kinetics of RYMV polyprotein maturation are still unknown. However, the proximity of the virulence sites to the triplet WAD which is conserved in VPg sobemoviruses [11] suggested that they belong to the putative VPg domain. Moreover, VPgs are involved in potyvirus virulence

towards several recessive resistances [see 2 and references cited]. Virulence mutations were generally localised in the central helix domain of the VPg [15]. Recessive resistance towards potyviruses were more often encoded by eukaryotic translation initiation factors 4E (eIF4E or eIF(iso)4E genes) [14]. Direct interactions involving *Potyviridae* VPgs and members of eIF4E family were demonstrated by several authors [8]. Members of the eIF4G family, another family of factors of the eIF complex that interact with eIF4Es, are also recruited by potyviruses via an undetermined (direct or indirect) mechanism [12]. In rice, the recessive resistance gene *rymv1* towards RYMV was identified and encodes eIF(iso)4G [1]. Several alleles of resistance namely *rymv1-2*, *rymv1-4* and *rymv1-3* were characterized by single substitutions or a short deletion, respectively [1]. The interaction mechanisms involved in the rice/RYMV pathosystem are still unknown. In this study, we determined the RYMV VPg termini, predicted structural domains and localised virulence sites. The 3D-topology and the biochemical properties of virulence mutations both suggest a direct site-to-site interaction between RYMV VPg and rice eIF(iso)4G.

RYMV VPg was isolated from purified virus particles using the protocol previously described [18]. Briefly, virions were purified from infected plants [3]. The genomic RNA was then extracted from 3 mg of purified particles using the RNeasy Plant RNA Kit (Qiagen) according to the manufacturer's instructions. Subsequently, the viral RNA linked to the VPg was hydrolyzed with 20% trifluoroacetic acid (TFA) for 48 h at room temperature. The purified VPg was blotted onto a PVDF membrane using a ProSorb cartridge with 300 μ l TFA 0.1% and rinsed twice with water. The PVDF-immobilized protein was directly subjected to N-terminal amino acid sequencing by automated Edman degradation (Applied Biosystems Procise Sequencer). Its N-terminal sequence was determined to be **XPFEIYXKFX**. Amino acids represented by X were undetermined. The residues identified represent amino acids 328-335 of the polyprotein P2a encoded by ORF2a and are located downstream of the protease domain as in other sobemoviruses (Figure 1). Based on the nucleotide sequence, the undetermined residues should be serine, glycine and arginine, respectively. The cleavage specificity of the protease was identified as E/(T,S) for *Sesbania mosaic virus* (SeMV), another sobemovirus [17]. Nine putative cleavage sites E/(T,S) were found in the RYMV polyprotein P2a (Figure 1). The N-terminal cleavage of VPg was deduced to be E³²⁶-S³²⁷. In the genus *Sobemovirus*, only the SeMV VPg C-terminal extremity was experimentally determined 77 amino acids downstream. By size homology, the C-terminal cleavage site of RYMV VPg is most likely E⁴⁰⁵-T⁴⁰⁶, producing a 79 amino acid-long protein from position 327 to 405 in the polyprotein P2a (Figure 1).

The amino acid composition of the RYMV VPg is characterised by a high proportion (39%) of charged residues (Figure 1). A total of 11 basic and 20 acidic amino acids are distributed along the protein. The secondary structure of the RYMV VPg was predicted using the server NPS@ (<http://npsa-pbil.ibcp.fr>)[4]. The consensus secondary structure prediction indicated at least two α -helices at positions 15-22 and 45-57 and two small β -strands at positions 26-30 and 37-39 (Figure 1). The predicted α -helices cover 35% of the VPg sequence. The analysis of twenty VPg sequences representative of the RYMV genetic diversity showed that the predicted secondary structures were conserved in all strains (data not shown).

The residue involved in virulence towards *rymv1-2*-mediated resistant rice reported earlier in position 374 [6] in P2a is located in position 48 within the VPg. This site is predicted to be in the central α -helix (Figure 1). In avirulent isolates, site 48 of VPg is occupied by a conserved arginine (R, basic), whereas this site is polymorphic (I, G, E, W) in virulent isolates [6, 13]. Glutamic acid (E) occupies site 48 in most resistance-breaking isolates. *Rymv1-2*-mediated resistance is characterized by a substitution from glutamic acid in eIF(iso)4G which is conserved in susceptible hosts to lysine (K, basic) at position 309 [1]. These opposite charge inversions R48E and E309K in the two proteins VPg and eIF(iso)4G respectively, are in accordance with a direct site-to-site interaction. VPg E48 occurred after two successive mutations via a glycine (G) step (mutational pathway I). Two other mutational pathways (II and III) which are isolate-specific were also observed. They involved either two successive mutations to valine (V) via an isoleucine (I) step, or one mutation to tryptophane (W). Residue-residue contact index preferences were derived from a set of 621 protein-protein interfaces of known high-resolution structure [5]. The pair R-K, two basic residues, had the lowest contact index, while pairs E-K or R-E with residues of opposite charges, had the highest contact indices (Figure 2). The stepwise fitness optimization process to achieve virulence along any of the three mutational pathways I (R/G/E), II (R/I/V) and III (R/W) at codon 48 paralleled an increase in the contact indices (Figure 2). This suggests an increase of the affinity with K309 of eIF(iso)4G throughout the successive amino-acids replacements. The optimal contact index was E x K, at a value similar to R x E (Figure 2). According to our hypothesis, E309 of susceptible eIF(iso)4G would interact with R48 of wild-type VPg and K309 of resistant eIF(iso)4G would interact with E48 of virulent VPg.

In addition to the major mutation at site 48 (80% of the virulence mutations), we recently found a few cases with other substitutions responsible for virulence in *rymv1-2* plants

[13]. The second best case (12% of the virulence mutations) was located in the same central helix of VPg at position 52. Interestingly, the sites 48 and 52 were predicted to locate along the longitudinal axis of the α -helix on the same side (Figure 3). Competition between variants at sites 48 and 52 always led to fixation of variant E48 in the viral population and displacement of the alternative virulent variants [13]. The competition and the close 3D-location of the sites 48 and 52 suggest they both participate in the same interaction domain. The mutation at position 52 which is a histidine-to-tyrosine substitution (H52Y) will also result in increasing the contact indices with K, but at a suboptimal level in comparison to the mutation 48E (Figure 2). Again, this could explain the displacement of the alternative virulent variants, in this case from the mutation 52 to the optimal mutation in codon 48. It is noteworthy that the site adjacent to 48 namely site 49 is under strong diversifying selection and was suggested to be involved in modulating the genetic barriers to virulence [13]. Indeed, the failure of the strain S2/S3 to follow mutational pathway I was correlated with a threonine at position 49. T49 was associated with K50. The amino acids adjacent to residue 48 may produce conformational changes that affect exposure of the site and thus alter the interaction. The influence of the context of an amino acid motif on its biological function was previously demonstrated for the DAG motif of potyvirus coat proteins on the aphid transmissibility [10]. The other minor virulence sites were localised in the loop upstream the central α -helix suggesting they might be included in the interaction domain.

The structure of the middle domain of eIF(iso)4G was modelled by homology with the human eIF4GII [1]. It suggested that the substitution E309K is located in an α -helix accessible at the surface of the protein and opposite from the interaction domain with eIF4A. Moreover, the resistance allele *rymv1-4* of *Oryza glaberrima* cultivar Tog5672 also harbours a single non-synonymous substitution from glutamic acid to lysine but at position 321 [1]. Although the substitution E321K was located in another α -helix, the two resistance mutations assembled in the same small surface patch suggesting they both belong to the same interaction domain. To further compare the role of E309K and E321K mutations, two virulent isolates with mutations E48 or Y52 were inoculated to the two resistant cultivars (20 plants for each cultivar). While all *rymv1-2* plants showed generalized systemic symptoms and high virus titer, none of the *rymv1-4* plant was infected. The mutations involved in virulence towards *rymv1-4*-mediated resistance are still unknown. Altogether, we propose a direct site-to-site interaction between one side of the central helix of RYMV VPg and one side of the central hairpin of eIF(iso)4G (Figure 3). The interaction domain of eIF(iso)4G with RYMV VPg is

far from the domain interacting with eIF4A, and resistance substitutions likely do not affect the functions of eIF(iso)4G in the host. By contrast, the interaction domain of eIF4E with potyvirus VPg was reported to partially overlap the cap binding domain and to eventually modify its function [19].

Translation initiation factors represent crucial partners for plant RNA viruses; these host factors have to be reached to accomplish the viral cycle [14]. This function can be performed by a dedicated protein, VPg (for review [16]). Data so far always demonstrated direct interactions involving *Potyviridae* VPgs (generally their central helix domains) and members of eIF4E family (eIF4E or eIF(iso)4E). Here we located the major virulence sites on the same side of the putative central helix of RYMV VPg. This location is compatible with a direct interaction with the middle domain of the rice eIF(iso)4G. Moreover, the opposite charge inversions in the two proteins and the stepwise increase of the contact indices are also in accordance with a direct interaction. In this study, we provided a model of direct site-to-site interaction between RYMV VPg and rice eIF(iso)4G, another way for a plant virus to interact with the translation initiation complex.

ACKNOWLEDGEMENTS

We thank Olivier Legall and Benoit Moury for helpful discussions and Anne-Lise Haenni for constructive criticisms of the manuscript.

REFERENCES

1. Albar L, Bangratz-Reyser M, Hebrard E, Ndjioudjop M-N, Jones M, Ghesquiere A (2006) Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to *Rice yellow mottle virus*. *Plant Journal* 47: 417-426
2. Ayme V, Souche S, Caranta C, Jacquemond M, Chadoeuf J, Palloix A, Moury B (2006) Different mutations in the genome-linked protein VPg of potato virus Y confer virulence on the pvr2(3) resistance in pepper. *Molecular Plant Microbe Interactions* 19: 557-563
3. Brugidou C, Opalka N, Yeager M, Beachy RN, Fauquet C (2002) Stability of rice yellow mottle virus and cellular compartmentalization during the infection process in *Oryza sativa* (L.). *Virology* 297: 98-108
4. Deleage G, Blanchet C, Geourjon C (1997) Protein structure prediction. Implications for the biologist. *Biochimie* 79: 681-686
5. Glaser F, Steinberg DM, Vakser IA, Ben-Tal N (2001) Residue frequencies and pairing preferences at protein-protein interfaces. *Proteins* 43: 89-102
6. Hébrard E, Pinel-Galzi A, Bersoult A, Siré C, Fargette D (2006) Emergence of a resistance-breaking isolate of *Rice yellow mottle virus* during serial inoculations is due to a single substitution in the genome-linked viral protein VPg. *Journal of General Virology* 87: 1369-1373

7. Hull R, Fargette D (2005) Sobemovirus. In: Fauquet C, Mayo MA, Maniloff J, Desselberger U and Ball LA (eds) *Virus Taxonomy*. Eight Report of the International Committee on Taxonomy of Viruses. Academic Press, Elsevier, pp 885-890
8. Khan MA, Miyoshi H, Ray S, Natsuaki T, Suehiro N, Goss DJ (2006) Interaction of genome-linked protein (VPg) of *Turnip mosaic virus* (TuMV) with wheat germ translation initiation factors eIF4E and eIF4F. *J. Biol. Chem.* 280: 28002-28010
9. Kouassi NK, N'Guessan P, Albar L, Fauquet C, Brugidou C (2005) Distribution and characterization of *Rice yellow mottle virus*: a threat to african farmers. *Plant Disease* 89: 124-133
10. Lopez-Moya JJ, Wang RY, Pirone TP (1999) Context of the coat protein DAG motif affects potyvirus transmissibility by aphids. *J Gen Virol* 80: 3281-3288
11. Makinen K, Tamm T, Naess V, Truve E, Puurand U, Munthe T, Saarma M (1995) Characterization of *Cocksfoot mottle sobemovirus* genomic RNA and sequence comparison with related viruses. *Journal of General Virology* 76: 2817-2825
12. Nicaise V, Gallois JL, Chafiai F, Allen LM, Schurdi-Levraud V, Browning KS, Candresse T, Caranta C, Le Gall O, German-Retana S (2007) Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in *Arabidopsis thaliana*. *FEBS Lett* 581: 1041-1046
13. Pinel-Galzi A, Rakotomalala M, Sangu E, Sorho F, Kanyeka Z, Traoré O, Sérémé D, Poulicard N, Rabenantaondro Y, Séré Y, Konaté G, Ghesquière A, Hébrard E, Fargette D (2007) Theme and variations in the evolutionary pathways to virulence of an RNA plant virus species. *PLoS Pathogens* 3: available on line
14. Robaglia C, Caranta C (2006) Translation initiation factors: a weak link in plant RNA virus infection. *Trends in Plant Science* 11: 40-45
15. Roudet-Tavert G, Michon T, Walter J, Delaunay T, Redondo E, Le Gall O (2007) Central domain of a potyvirus VPg is involved in the interaction with the host translation initiation factor eIF4E and the viral protein HcPro. *J Gen Virol* 88: 1029-1033
16. Sadowy E, Milner M, Haenni AL (2001) Proteins attached to viral genomes are multifunctional. *Adv Virus Res* 57: 185-262
17. Satheskumar PS, Lokesh GL, Savithri HS (2004) Polyprotein processing: cis and trans proteolytic activities of *Sesbania mosaic virus* serine protease. *Virology* 318: 429-438
18. van der Wilk F, Verbeek M, Dullemans AM, van den Heuvel JF (1997) The genome-linked protein of *Potato leafroll virus* is located downstream of the putative protease domain of the ORF1 product. *Virology* 234: 300-303
19. Yeam I, Cavatorta JR, Ripoll DR, Kang BC, Jahn MM (2007) Functional dissection of naturally occurring amino acid substitutions in eIF4E that confers recessive potyvirus resistance in plants. *Plant Cell* 19: 2913-2928

Figure 1: Genomic organization of RYMV, location and sequence analysis of the VPg.

Top. The N-terminus of the VPg was sequenced and located at serine 327 in the polyprotein P2a. Among the nine putative cleavage sites (arrows), the VPg C-terminal site is predicted to be E405/T406. VPg (grey) is encoded by ORF2a downstream of the protease and upstream of the frameshifting signal (fs). Bottom. VPg is a 79 amino acid-long protein. Acidic (red) and basic (blue) residues, and consensus predicted α -helices and β -strands are indicated. The major 48 and the minor 52 sites (triangles) are involved in RYMV virulence on the highly resistant rice cultivar Gigante. Site 49 (asterisk) was reported to be under diversifying selection. The WAD motif (box) is conserved in sobemoviruses.

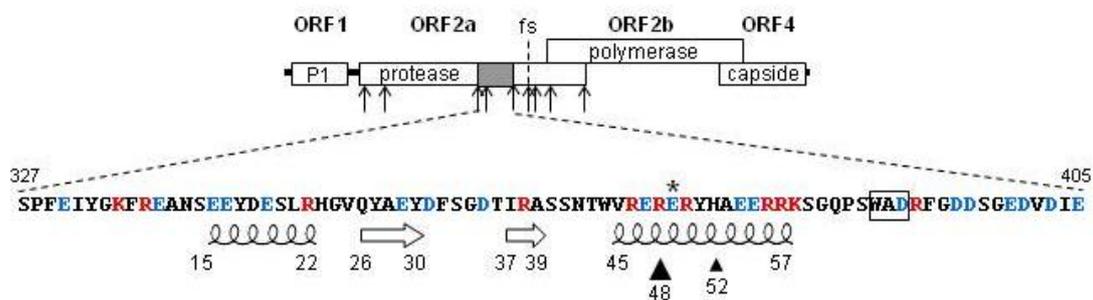


Figure 2: Contact indice increase of RYMV virulent isolates. Left, contact indice between glutamic acid (E) and arginine (R) as previously evaluated (Glaser et al, 2001). The susceptible allele *rymv1-1* harbours E at position 309 of eIF(iso)4G and wild-type virus harbours R at position 48 of VPg. Right, ranked contact indices between lysine (K) and the 21 amino acids. Resistance allele *rymv1-2* is mutated in lysine and virulent VPg is polymorphic at position 48 (Pinel-Galzi et al., 2007). Three mutational pathways to RYMV virulence were observed in *rymv1-2*-mediated resistant plants. The major pathway I (red) involves two successive mutations from R to E via a transient glycine (G). The pathways II (blue) and III (green) are isolate-specific and involve either mutations from R to isoleucine (I) and sometimes to valine (V), or mutation from R to tryptophane (W), respectively. All mutational pathways resulted in increased contact indices with K.

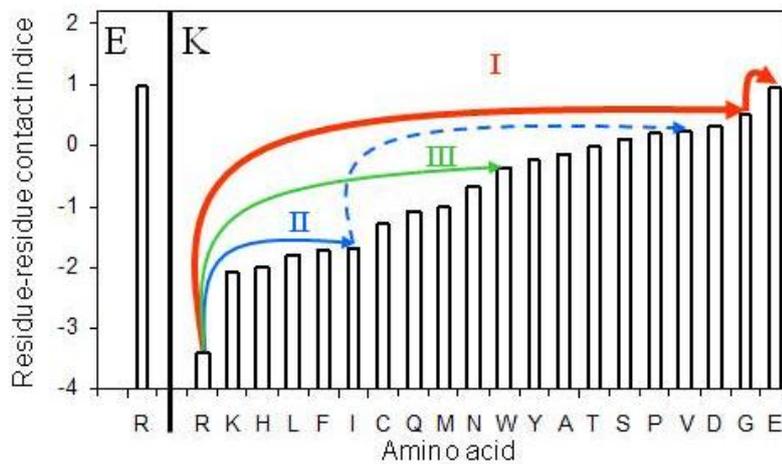


Figure 3: Hypothetical model of the VPg/eIF(iso)4G interaction. At least one side of the central α -helix of VPg is proposed to interact with one side of the α -hairpin of the middle domain of eIF(iso)4G. R48 of VPg is predicted to play a major role in the interaction with E309 of eIF(iso)4G. H52 located in the same side of the VPg helix would belong to the interaction domain. The orientation of the VPg central helix compared to the two antiparallel helices forming a hairpin in the eIF(iso)4G middle domain remain to be determined. Site 321 mutated in the *rymv1-4* allele and located on the same side of the hairpin of eIF(iso)4G belongs to the interaction domain with VPg. The mutations E48 and Y52 did not induce virulence against this allele. The VPg sites interacting with E321 are still unknown.

