Published in final edited form as:

Biomol NMR Assign. 2019 April ; 13(1): 9–13. doi:10.1007/s12104-018-9842-3.

Chemical shift assignment of the viral protein genome-linked (VPg) from Potato virus Y

Luciana Coutinho de Oliveira, Laurent Volpon, Michael J. Osborne, and Katherine L.B. Borden*

Institute of Research in Immunology and Cancer (IRIC), Department of Pathology and Cell Biology, Université de Montréal, Pavillon Marcelle-Coutu, Chemin Polytechnique, Montreal, Qc, Canada.

Abstract

The dysregulation of translation contributes to many pathogenic conditions in humans. Discovering new translational mechanisms is important to understanding the diversity of this process and its potential mechanisms. Such mechanisms can be initially observed in viruses. With this in mind, we studied the viral protein genome-linked VPg factor from the largest genus of plant viruses. Studies in plants show that VPg binds to the eukaryotic translation initiation factor eIF4E for translation of viral RNAs. VPg contains no known eIF4E binding motifs and no sequence homology to any known proteins. Thus, as a first step in understanding the structural basis of this interaction, we carried out NMR assignments of the VPg from the potato virus Y (PVY) potyvirus protein.

Keywords

VPg; Potyviruses; translation; eIF4E; cancer

Biological context

Potyvirus form the largest genus of plant viruses and affect major agricultural crops such as potatoes and lettuce (Bastet et al. 2017; Robaglia and Caranta 2006). These viruses are single-stranded positive-sense RNA viruses in the picorna-like super group of viruses (Bastet et al. 2017; Robaglia and Caranta 2006). In these studies, there is an interesting array of biochemistry not the least of which is that one viral factor, the viral protein genome-linked (VPg) protein, is covalently attached to the 5' end of the genomic RNA. This physical association is required for translation of the genomic RNA and the viral life cycle. Plant viruses only encode a few proteins (between 4–10) and thus heavily rely on host cell factors for their replication (Robaglia and Caranta 2006). In particular, potyviruses have demonstrated a strong reliance on the host cell translation machinery. Thus, many of the naturally occurring resistance mechanisms in plants are derived from mutation of these factors (Bastet et al. 2017; German-Retana et al. 2008; Robaglia and Caranta 2006). For instance, one of the major naturally occurring forms of resistance for potyviruses arises due

*To whom correspondence should be addressed. katherine.borden@umontreal.ca.

to mutations in the host plant translation initiation factor eIF4E (Bastet et al. 2017; German-Retana et al. 2008; Robaglia and Caranta 2006). This naturally occurring set of mutations provides strong evidence for the requirement of eIF4E. Indeed, the link is strong enough that the idea has emerged that cultivars should be developed with eIF4E mutations to improve crop yields (Bastet et al. 2017; Wang and Krishnaswamy 2012). Given the association of VPg with the viral RNA, it has been postulated that VPg directly binds eIF4E in order to translate the viral RNA and continue the life cycle. It has been further hypothesized that such an association suppresses host cell translation through sequestration of eIF4E.

eIF4E is dysregulated in many malignancies in humans where its elevation is linked to poor prognosis (Carroll and Borden 2013). Viruses often provide unique molecular insights, which later turn out to be relevant for mammalian cell biology. The interaction of VPg and eIF4E in plants can thereby serve as a model to elucidate novel regulatory processes in humans. Indeed, plant and human eIF4E are highly conserved both at the sequence and structural level. Thus, VPg could provide a unique tool to study novel ways to regulate eIF4E activity in both plants and animals. Consistently, VPgs do not contain any known eIF4E binding motifs. For instance, some eIF4E binding proteins contain a consensus binding site comprised of a YXXXXLΦ (where X is any residue and Φ is any hydrophobic) as observed in eIF4G, eIF4E binding proteins (4E-BPs) and many homeodomains; others use RING domains such as in promyelocytic leukemia protein PML, the arenavirus protein Z, and the human homologue of ariadne HHARI; and yet others use GIGYF1/2 domains (Carroll and Borden 2013; Kentsis et al. 2001; Osborne and Borden 2015; Peter et al. 2017; Tan et al. 2003; Topisirovic and Borden 2005; Volpon et al. 2010). Sequence analysis reveal that no such domains exist in the potyvirus VPgs and further, they have no sequence homology to other proteins in the database including VPgs from other virus families. There have been no structural studies on any potyvirus proteins reported. As a first step to gain insights into the mechanisms underpinning the activities of VPgs and their relationship to eIF4E, we expressed and completed the NMR assignments of the potyvirus potato virus Y (PVY) VPg protein.

Methods and experiments

Protein expression and purification

The sequence encoding Viral Protein genome-linked from Potato virus Y (PVY) was chemically synthesized with codon optimization for expression in *E. coli* (GenScript) and cloned into the expression vector pET-28a between the BamHI and XhoI restriction sites from residues 1 to 188 and 38 to 188. In addition, a TEV protease cleavage site was inserted between the T7 tag and VPg leaving the sequence Gly-Ser attached to the N-terminus of the protein. The plasmid was overexpressed in *Escherichia coli* BL21(DE3) cells in minimum media (see below) at 37°C and induced overnight by 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 20°C. Cells were harvested by centrifugation and stored at -80°C until use. The frozen cells were resuspended in lysis buffer consisting of 50mM Tris pH 8.0, 1M NaCl, 7mM β-mercaptoethanol (βME), 10mM imidazole, 0.5% NP40 and cocktail of protease inhibitor (GE). The reducing agent was used in all buffers to prevent any possible formation of disulfides from the cysteine residue. After the cells were lysed by

sonication, lysozyme from chicken egg white (Sigma) was added at a final concentration of 1 mg/mL and cells were mixed at 4°C for 15 min. The lysate was cleared by centrifugation (30 min, 20,000 rpm, 4°C) and purified over Ni-NTA beads (Qiagen) onto a gravity flow column. After the resin was extensively washed in 50mM Tris pH 8.0, 1M NaCl, 7mM β -mercaptoethanol (β ME), 20mM imidazole, the protein was eluted in the same buffer but containing 500mM imidazole. Then, the TEV protease was added to VPg to remove both the N-terminal His tag and T7 tag, and was dialyzed overnight against PBS buffer containing 10mM EDTA and 1mM DTT. High level of purity (>95%) was achieved by gel filtration on Superdex 75pg column (GE-Biosciences) in 50mM Sodium phosphate (pH 7.5), 150mM NaCl, 1mM DTT.

Isotope labeling schemes

Double ($^{15}\text{N}/^{13}\text{C}$) and triple ($^2\text{H}/^{15}\text{N}/^{13}\text{C}$) labeled VPg was prepared from cells grown on minimal M9 media containing 2g/L [^{15}N]ammonium chloride and 2g/L [$^{13}\text{C}_6$] or [$^{13}\text{C}_6$, ^2H]glucose (Sigma-Aldrich). For the deuterated VPg, *E. coli* strains were adapted to deuterated minimal medium by gradually increasing the deuterium content. Briefly, freshly transformed colonies were used to start a 25ml culture in LB medium/ H_2O , followed by successive pre-cultures in M9 media containing 0%, 50% and 100% D_2O . Each pre-culture was initiated to an OD600 of 0.25, and the final bacterial culture was overexpressed for 24 h.

NMR spectroscopy and data processing

The solution conditions used for assignment of VPg consisted of 0.4–0.5 mM protein in 50 mM phosphate buffer (pH 7.5), 150 mM NaCl, 1 mM DTT, 0.02% NaN_3 , with either 7 % or 100 % D_2O . For the latter, VPg was exchanged in completely deuterated buffer using Millipore Amicon Centrifugal Filter (10 kDa cut-off). VPg was transferred to an NMR Shigemi tube and NMR experiments were acquired at 20°C on a Bruker AVANCE III HD 600 spectrometer with a QCIP Z-axis gradient cryoprobe, or on a Bruker 800 MHz spectrometer equipped with TCI cryogenic probe. The experiments acquired in 93 % $\text{H}_2\text{O}/7\%$ D_2O are the 2D ^1H - ^{15}N HSQC and the 3D HNCA, HN(CO)CA, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, CC(CO)NH and H(CC)(CO)NH. In addition, a backbone $^{15}\text{N}\{-^1\text{H}\}$ heteronuclear NOE (het-NOE) experiment was performed at 800 MHz in interleaved fashion using saturation time and inter-scan relaxation delay of 3 sec. The steady-state het-NOE values were obtained from the cross peak intensity ratios of spectra recorded starting from the pure ^{15}N polarization with or without presaturation of the ^1H spins. Two ^1H - ^{13}C HSQCs (aliphatic and aromatic regions) together with a 3D HCCH-TOCSY and an aromatic-NOESY were obtained in 100% D_2O . Most of the 3D spectra (except for side chain assignment) were acquired in a non-uniform manner using the Bruker standard parameter sequences with 15 to 20% Poisson Gap Sampling Schemes (Hyberts et al. 2010). Non-uniform sampling spectra were processed using SMILE and NMRPipe (Delaglio et al. 1995) and analyzed with CcpNmr Analysis (Vranken et al. 2005) and Sparky (Goddard and Kneller 2003).

Extent of assignments and data deposition

VPg consists of 188 residues but the full-length construct had high tendency to precipitate and exhibited a poor quality HSQC spectra with broad signals. Assigning full-length VPg

would be challenging, thus progressive N-terminal deletions were generated and a stable construct consisting of residues 38 to 188 was obtained (noted VPg 37). The ^1H - ^{15}N HSQC spectrum of VPg 37 is shown in Fig. 1a and exhibited large chemical shift dispersion for most of the signals, indicative of a well-ordered three dimensional protein structure. We also noticed the presence of strong cross-peaks in the center of the spectra, characteristic of intrinsically disordered regions.

Using the set of 3D experiments listed above, we were able to assign unambiguously 94 % of the ^{15}N and the $^1\text{H}^{\text{N}}$ of non-proline residues of VPg 37 in the region 70 to 188 (missing S120, R129-W132, H144, N145, G170; Fig. 1b, residues in yellow). In addition, 87% of the $^{13}\text{C}'$, 92% of the $\text{C}\alpha$, 93% of the $\text{C}\beta$, 76% of the $\text{H}\alpha$ and 52% of the $\text{H}\beta$ resonances were assigned. In total, completeness of ^1H resonances assignment is 78.4%, including side chains. Regarding the N-terminal portion of this construct (residues 38–69), only two short segments could be assigned (Y40-G44 and T49-G52). This could be explained by the presence of GK repeats in the N-terminal end (Fig. 1b) and by peak broadening due to the intermediate exchange regime or to the high exchange rate observed at basic pH. Indeed, by lowering the pH gradually to 6.0, while a few peaks disappeared, we also observed the appearance of new peaks, especially in the glycine region probably coming from the glycines of the N-terminal region (residues 38–69). The chemical shift values for ^1H , ^{13}C and ^{15}N resonances have been deposited into the BioMagResBank under accession number 27506.

Secondary structure elements of VPg 37 were identified by the TALOS-N server (Shen and Bax 2013) using C' , $\text{C}\alpha$, $\text{C}\beta$, $\text{H}\alpha$ and N chemical shifts (Fig. 2a). This analysis revealed the presence of five β -strands (74–77; 84–87; 123–127; 135–142; 173–176) and three α -helices (93–109; 115–118; 177–178) arranged in the sequential order $\beta 1$ – $\beta 2$ – $\alpha 1$ – $\alpha 2$ – $\beta 3$ – $\beta 4$ – $\beta 5$ – $\alpha 3$. In addition, we have carried out ^{15}N - $\{^1\text{H}\}$ heteronuclear relaxation NMR experiment to further characterize conformational dynamics at the pico- to nanosecond timescale (Fig. 2b). The secondary structures of VPg 37 are supported by this experiment since N-H bond vectors that undergo motion faster than the overall tumbling of the molecule show a decreased NOE intensity relative to the structure elements. In particular, the few residues assigned in the N-terminal end (see above), the loop $\beta 4$ – $\beta 5$ and the C-terminal end (182–188) showed lower values compared to the rest of the protein. We note that the N-terminus is not completely unstructured as shown by values ranging from 0.3 and 0.6. This could be due to fast conformational exchange or to the presence of a transitory state in which the N-terminus interacts with the folded region of VPg. Finally, as can be seen in the ^1H - ^{15}N HSQC spectrum (Fig. 1a), a few cross peaks could not be assigned. These could arise from the N-terminal region but also from the long flexible loop between $\beta 4$ – $\beta 5$ in which it was identified the presence of a second resonance assignment pathway for example for the residues 156–157 and 163–165 (shown in orange in Fig 1a and 1b). This could arise from a cis/trans isomerization of P162. Altogether, this suggests the presence of a rigid structure constituted mainly by a β -sheet and a long α -helix, together with a long and flexible loop. In all, this provides a first step in the determination of how VPg associates with eIF4E.

Acknowledgments.

800 MHz NMR data were acquired at QANUC and 600 MHz NMR data at IRIC which were both supported in part by CFI. This work was supported by the following grants to KLBB: RO1 NIH 98571; RO1 NIH80728 and KLBB holds the Canada Research Chair in Molecular Biology of the Cell Nucleus.

References

- Bastet A, Robaglia C, Gallois J (2017) eIF4E Resistance: Natural Variation should guide gene editing *TRENDS in Plant Science* 22:411–419 [PubMed: 28258958]
- Carroll M, Borden KL (2013) The oncogene eIF4E: using biochemical insights to target cancer *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* 33:227–238
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes [see comments] *J Biomol NMR* 6:277–293 [PubMed: 8520220]
- German-Retana S et al. (2008) Mutational analysis of plant cap-binding protein eIF4E reveals key amino acids involved in biochemical functions and potyvirus infection *J Virol* 82:7601–7612 [PubMed: 18480444]
- Goddard TD, Kneller DG (2003) Sparky 3, University of California, San Francisco, CA
- Hyberts SG, Takeuchi K, Wagner G (2010) Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data *J Am Chem Soc* 132:2145–2147 [PubMed: 20121194]
- Kentsis A, Dwyer EC, Perez JM, Sharma M, Chen A, Pan ZQ, Borden KL (2001) The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E *J Mol Biol* 312:609–623 [PubMed: 11575918]
- Osborne MJ, Borden KL (2015) The eukaryotic translation initiation factor eIF4E in the nucleus: taking the road less traveled *Immunol Rev* 263:210–223 [PubMed: 25510279]
- Peter D et al. (2017) GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression *Genes Dev* 31:1147–1161 [PubMed: 28698298]
- Robaglia C, Caranta C (2006) Translation initiation factors: a weak link in plant virus infection *TRENDS in Plant Science* 11:40–45 [PubMed: 16343979]
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks *J Biomol NMR* 56:227–241 [PubMed: 23728592]
- Tan NG, Ardley HC, Scott GB, Rose SA, Markham AF, Robinson PA (2003) Human homologue of ariadne promotes the ubiquitylation of translation initiation factor 4E homologous protein, 4EHP *FEBS letters* 554:501–504 [PubMed: 14623119]
- Topisirovic I, Borden KL (2005) Homeodomain proteins and eukaryotic translation initiation factor 4E (eIF4E): an unexpected relationship *Histol Histopathol* 20:1275–1284 [PubMed: 16136508]
- Volpon L, Osborne MJ, Capul AA, de la Torre JC, Borden KL (2010) Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E *Proceedings of the National Academy of Sciences of the United States of America* 107:5441–5446 [PubMed: 20212144]
- Vranken WF et al. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline *Proteins* 59:687–696 [PubMed: 15815974]
- Wang A, Krishnaswamy S (2012) Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement: eIF4E mediated resistance to plant viruses *Molecular Plant Pathology* 13:795–803 [PubMed: 22379950]

Fig. 1. (a) ^1H - ^{15}N HSQC spectrum of 400 μM VPg 37 in 50 mM sodium phosphate, 150 mM NaCl, 1 mM DTT (pH 7.5) at 293 K. Assignment of the backbone amide resonances (black) and side chain amides of Asn/Gln residues (blue) are shown. (b) Amino acid sequence of VPg 37. The first two residues (shown in lowercase) are from the tag. Amino acids for which the backbone was assigned are highlighted in yellow. Sequences showing two states of resonances are in orange (the second state for each residue is shown in panel 1a with an orange line), and Prolines are in blue.

Fig. 2.

(a) Secondary structure prediction of VPg 37 from TALOS-N derived from ^{15}N , $^{13}\text{C}'$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ and $\text{H}\alpha$ chemical shifts. The probability of β -strand (blue) or α -helical (red) structure is represented by the height of the bars as predicted by the TALOS-N software. The expected secondary structure elements are represented above. (b) Plot of backbone heteronuclear ^{15}N - $\{^1\text{H}\}$ NOE values versus residue number for VPg 37. Values missing are either Prolines, unassigned residues due to overlap or absent in the HSQC.