



Article

VPg Impact on Ryegrass Mottle Virus Serine-like 3C Protease Proteolysis and Structure

Gints Kalnins ^{1,†}, Rebeka Ludviga ^{2,†}, Ieva Kalnciema ³, Gunta Resevica ³, Vilija Zeltina ³, Janis Bogans ⁴, Kaspars Tars ¹, Andris Zeltins ³ and Ina Balke ^{2,*} 

¹ Structural Biology Group, Latvian Biomedical Research and Study Centre, Ratsupites Street 1, k-1, LV-1067 Riga, Latvia

² Plant Virus Protein Research Group, Latvian Biomedical Research and Study Centre, Ratsupites Street 1, k-1, LV-1067 Riga, Latvia

³ Plant Virology Group, Latvian Biomedical Research and Study Centre, Ratsupites Street 1, k-1, LV-1067 Riga, Latvia

⁴ Biotechnology Core Facility, Latvian Biomedical Research and Study Centre, Ratsupites Street 1, k-1, LV-1067 Riga, Latvia

* Correspondence: inab@biomed.lu.lv

† These authors contributed equally to this work.

Abstract: Sobemoviruses encode serine-like 3C proteases (Pro) that participate in the processing and maturation of other virus-encoded proteins. Its *cis* and *trans* activity is mediated by the naturally unfolded virus-genome-linked protein (VPg). Nuclear magnetic resonance studies show a Pro–VPg complex interaction and VPg tertiary structure; however, information regarding structural changes of the Pro–VPg complex during interaction is lacking. Here, we solved a full Pro–VPg 3D structure of ryegrass mottle virus (RGMoV) that demonstrates the structural changes in three different conformations due to VPg interaction with Pro. We identified a unique site of VPg interaction with Pro that was not observed in other sobemoviruses, and observed different conformations of the Pro β 2 barrel. This is the first report of a full plant Pro crystal structure with its VPg cofactor. We also confirmed the existence of an unusual previously unmapped cleavage site for sobemovirus Pro in the transmembrane domain: E/A. We demonstrated that RGMoV Pro *in cis* activity is not regulated by VPg and that *in trans*, VPg can also mediate Pro in free form. Additionally, we observed Ca^{2+} and Zn^{2+} inhibitory effects on the Pro cleavage activity.

Keywords: ryegrass mottle virus; serine-like 3C proteases; VPg; *Sobemovirus*



Citation: Kalnins, G.; Ludviga, R.; Kalnciema, I.; Resevica, G.; Zeltina, V.; Bogans, J.; Tars, K.; Zeltins, A.; Balke, I. VPg Impact on Ryegrass Mottle Virus Serine-like 3C Protease Proteolysis and Structure. *Int. J. Mol. Sci.* **2023**, *24*, 5347. <https://doi.org/10.3390/ijms24065347>

Academic Editor: Irmgard Tegeder

Received: 7 February 2023

Revised: 3 March 2023

Accepted: 9 March 2023

Published: 10 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Ryegrass mottle virus (RGMoV) belongs to the family *Sobemoviridae*, genus *Sobemovirus*, and infects monocotyledons, such as Italian ryegrass (*Lolium multiflorum*), cocksfoot (*Dactylis glomerata*), wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), oats (*Avena sativa*), rye (*Secale cereale*), and Italian millet (*Setaria italica*) [1,2]. RGMoV was isolated from Italian ryegrass and cocksfoot in Japan [1] and was also found in irrigation waters in Ecuador in 2021 [3]. Its genome is a monocistronic single-stranded positive-sense RNA that is 4212 or 4247 nucleotides in length [3,4]. Sobemovirus genomic RNA (gRNA) encodes five open reading frames (ORF), where ORF1, ORFx, ORF2a, and ORF2b are translated from the gRNA, and ORF3, which encodes the coat protein (CP), is translated from a subgenomic RNA (sgRNA) [5,6]. gRNA and sgRNAs lack a cap structure at the 5'-end and a poly(A)-tail at the 3'-end [7]. Instead, the 5'-end is protected by covalently bound VPg (mimicking eukaryotic cap structure) in RGMoV, connecting the OH group of the first serine (S) and a terminal 5'-phosphate group of the RNA genome via a phosphodiester bond [8]. ORF2a and ORF2b of sobemoviruses encode two polyproteins, where the N-terminal parts of ORF2a and ORF2ab encode serine-like 3C protease (Pro), which displays a significant

similarity to the cysteine (C) proteases of picornaviruses [9]. The next domain encoded by both ORFs is VPg, followed by P18, which can be divided into two domains: RNA-binding (P10) and an ATPase domain (P8) in the C-terminal part of ORF2a [10]. The C-terminal part of polyprotein P2ab contains an RNA-dependent RNA polymerase motif (RdRp). In sobemoviruses, the location of VPg between Pro and RdRp is unique among phylogenetically related polero-, enamo-, and barnaviruses [11]. Functional domains from polyproteins of sobemoviruses are released by proteolytic processing by Pro [8,12,13]. Pro catalyzes the hydrolysis of peptide bonds located between two specific residues, the first of which is glutamic acid (E). Identification of sobemoviral VPgs attached to the viral genomes indicated that the 2a polyprotein was processed for southern bean mosaic virus (SBMV) at E/T (glutamic acid/threonine) residues at positions 325/326 and 402/403, cocksfoot mottle virus (CfMV) at E³¹⁷/N³¹⁸ (glutamic acid/asparagine) and E³⁹⁵/T³⁹⁶ residues, rice yellow mottle virus (RYMV) at E³²⁶/S³²⁷ and E⁴⁰⁴/T⁴⁰⁵ residues, and RGMoV at E/S residues at positions 314/315 and 393/394 [8,13]. VPg can directly regulate Pro activity, as Sesbania mosaic virus (SeMV) Pro is active *in trans* only in fusion with VPg [14]. Mutagenesis analysis revealed that residue W⁴³ (tryptophan) of VPg interacts with SeMV Pro residues W²⁷¹ and H²⁷⁵ (histidine), activating Pro [15]. VPgs of plant viruses are characterized as “natively unfolded proteins”, lacking a well-defined 3D structure and existing in multiple conformations at physiological conditions [16]. These proteins adopt a rigid conformation stabilized *in vivo* upon interaction with natural substrates and are multifunctional [11]. Bioinformatics methods have unfolded the nature of VPgs of several sobemoviruses (SeMV, RGMoV, CfMV, RYMV, SBMV, and southern cowpea mosaic virus (SCPMV)), potyviruses (lettuce mosaic virus (LMV), potato virus Y (PVY), potato virus A (PVA), tobacco etch virus, turnip mosaic virus, bean yellow mosaic virus), and *Caliciviridae* family members (rabbit hemorrhagic disease virus (genus *Lagovirus*), vesicular exanthema of swine virus (genus *Vesivirus*), sapporo virus Manchester virus (genus *Sapovirus*), and Norwalk virus (genus *Norovirus*)) [14,16–18]. In VPgs of the RYMV group (to which RGMoV belongs), similar to potyvirus and *Caliciviridae* family members, the N- and C-termini are predicted to be disordered [16,18]. Secondary structure prediction in this sobemovirus group indicated the presence of an α -helix, followed by two β -strands, and another α -helix. Some of the terminal regions of these VPgs are predicted to have propensities that are both disordered and folded into α -helices. Residues 48 and 52, which are associated with RYMV virulence, are located in the C-terminal region [19]. Recent nuclear magnetic resonance (NMR) studies of SeMV VPg have demonstrated that it has a novel tertiary structure with an α - β - β - β topology [20]. Circular dichroism (CD) spectrum analysis for SeMV, RYMV, LMV, PVY, and PVA also revealed features typical of intrinsically disordered proteins [14,16,18,21]. NMR analysis of VPgs was also performed for PVY, PVA, and cowpea mosaic virus [18,21,22]. PVA VPg behaves as a partially folded species that contains a hydrophobic core domain [21]. Currently, the 3D structures of VPg peptides are known only for viruses infecting mammals. Determination of the 3D structure of the first 15 amino acids (aa) of the N-terminus of foot-and-mouth disease virus (family *Picornaviridae*) VPg1 was possible only with RdRp, and the C-terminus remained completely disordered [23,24]. A similar structure was observed for enterovirus 71 (family *Picornaviridae*) VPg, where the first 20 aa formed a V-shaped extended conformation from the front side of the catalytic center to the back of RdRp [25]. The 3D structure of coxsackievirus B3 (family *Picornaviridae*) VPg 7–15 aa [26], the 3D structure of the feline calicivirus (FCV), and murine norovirus (MNV) VPg core of the protein adopts a compact helical structure flanked by flexible N- and C-termini. The core (10–76 aa) of FCV VPg (10–76 aa; PDB ID 2M4H) contains a well-defined three-helix bundle, and the MNV VPg core (11–85; PDB ID 2M4G) contains only the first two of these secondary structure elements [27]. The same structure remains during the interaction with the palm domain of the MNV RdRp [28].

Sobemovirus Pro is a glutamyl endopeptidase (GEPases). GEPases are serine proteases that preferentially cleave peptide bonds following E [29,30]. GEPases have been identified and characterized in gram-positive bacteria and (+)RNA viruses. All GEPases belong to

the chymotrypsin structural family, one of the most extensively studied enzyme families. Chymotrypsin-like protease (CLP) molecules share their spatial organization principle, the so-called chymotrypsin (or trypsin) fold [31], which comprises two β -barrels with an active site located between them. The only 3D structure of the sobemovirus Pro domain available in the worldwide Protein Data Bank (PDB) is that of SeMV Pro, determined at a resolution of 2.4 Å (PDB ID 1zyo). Its 3D structure has a typical CLP fold with a well-formed active site and a substrate-binding cleft that is more similar to that of cellular [in particular, Glu-specific protease from *Streptomyces griseus* (Glu-SGP; PDB ID 1 hpg)] rather than viral Pro [12,32]. Pro carries the conventional D-H-S catalytic triad, and modification of its residues terminates polyprotein processing [12]. The consensus aa sequence around the catalytic triad is H(X^{32–35})[D/E](X^{61–62})TXXGXSG [33]. Similarly, for all GEPases, the conserved residues of SeMV Pro H²¹³⁽²⁹⁸⁾ and T¹⁹⁰⁽²⁷⁹⁾ are maintained, located within the S1 site of the enzyme. However, position 216(301) is occupied by a large hydrophobic residue, phenylalanine (F; [31]). SeMV Pro comprises two β -barrels connected by a long interdomain loop. Both the active site and substrate-binding cleft are located between the two barrels and are fairly exposed to a solvent. Mutation analysis of the glutamate-binding site (S1-binding pocket) residues H²⁹⁸, T²⁷⁹, and N³⁰⁸ of SeMV Pro demonstrated that these are indeed crucial for Pro activity. In addition, several downstream residues are shown to be important for Pro activity [32]. The substrate specificities of SeMV Pro were predicted to be N, Q-E/T, and S-X (where X is an aliphatic residue) [10]. However, multiple sequence alignments revealed no common substrate specificity for all sobemoviruses, except that the S1 binding pocket seems to be highly specific for glutamate (Glu) or glutamine (Q) [34].

In this study, we demonstrated that a 10-residue-long VPg peptide in the C-terminal region (GEMTWADMVE) can assume a helical conformation and bind to a specific site in the Pro domain, causing a conformational change. We also confirmed an additional Pro cleavage site that has not been mapped for sobemoviruses after a transmembrane domain (TMD) between E/A (glutamic acid/alanine) residues, and demonstrated that *in trans* Pro activity can be induced by non-covalent protein–protein interactions with VPg, without the requirement of fusion to the Pro domain, as was reported previously [14].

2. Results

2.1. Description of Serine-like 3C Protease, Its Fusion Variant with Cofactor VPg, and the Catalytic Triad S Mutant Used for Structure Determination

An RGMoV cDNA clone (GenBank No. EF091714, [4]) was used as a DNA template to amplify the N-terminal truncated Pro coding domain (Δ 50Pro) lacking 50 residues of TMD (counting from the first methionine (M), predicted by the TMHMM program [35]), Δ 117Pro catalytic triad S¹⁵⁹ mutant (Δ 117Pro^{cm}, S159A), and Δ 50Pro fusion version with VPg (Δ 50Pro-E/A-VPg), where both predicted Pro cleavage sites between Pro and VPg domains were mutated—E³¹⁵ and E³¹⁸ to A (E315A and E318A)—to prevent VPg processing from Pro. The constructs contained an N-terminal 6xH-tag for purification by Ni²⁺-immobilized metal affinity chromatography (IMAC). After expression and protein purification on IMAC and sample analysis by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE), we observed that the molecular weights (MWs) of the Δ 50Pro and Δ 50Pro-E/A-VPg products were lower than expected. Instead of 29.5 kDa for Δ 50Pro, it was 20.9 kDa, and for Δ 50Pro-E/A-VPg, while the predicted MW was 38.5 kDa, the observed MW was 29.5 kDa (Figure 1B).

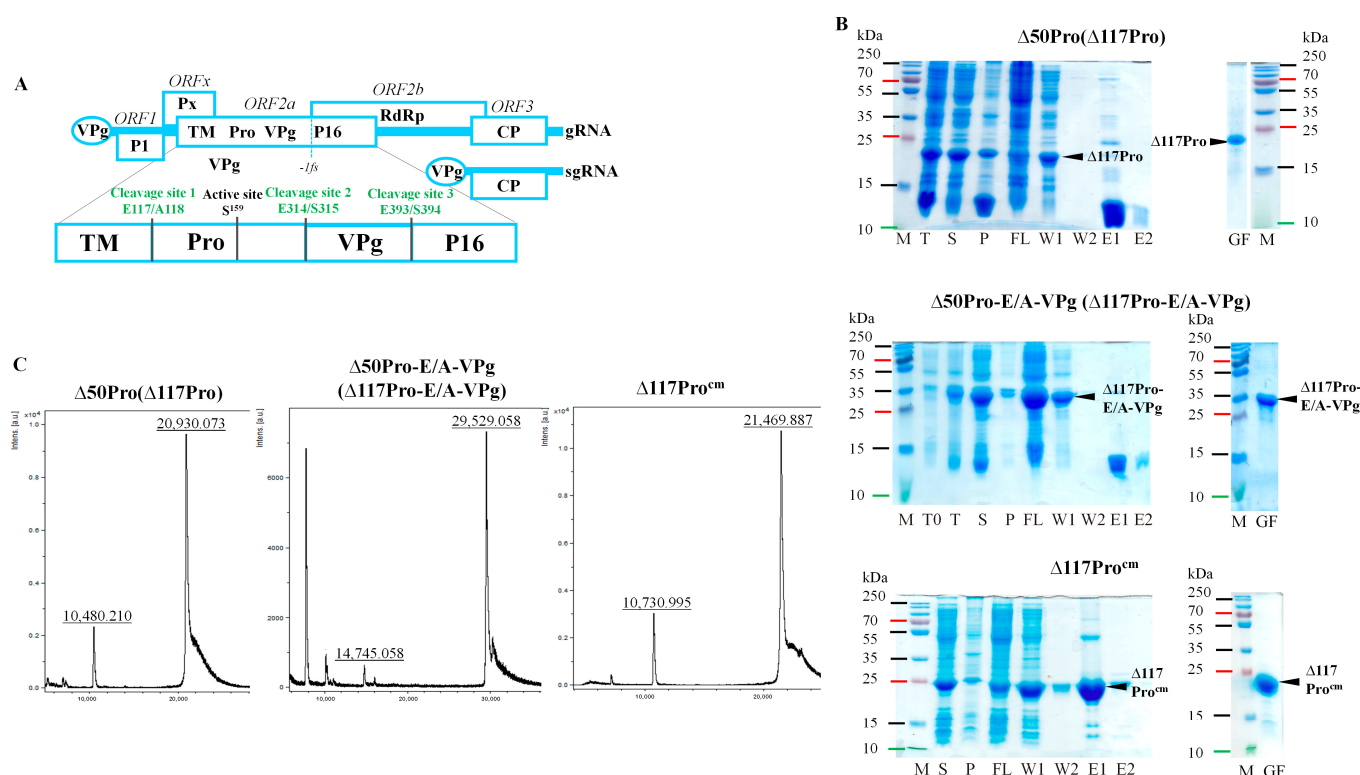


Figure 1. $\Delta 50\text{Pro}$, $\Delta 50\text{Pro-E/A-VPg}$, and $\Delta 117\text{Pro}^{\text{cm}}$ expression and analysis. (A)—RGMoV schematic overview; number at aa—location in ORF2a starting from the first M; TM: transmembrane domain; Pro: protease; VPg: virus-genome like protein; P16: protein with 16 kDa MW; RdRp: RNA dependent RNA polymerase; (B)—SDS-PAGE analysis of expression and purification of $\Delta 50\text{Pro}$, $\Delta 50\text{Pro-E/A-VPg}$, and $\Delta 117\text{Pro}^{\text{cm}}$; M—protein marker (Page Ruler Plus, Thermo Fisher Scientific, Waltham, MA, USA); T0—total cell lysate before expression; T—total cell lysate after expression; S—soluble protein fraction; P—pellet; FL—flow fraction after IMAC column; W1—first wash; W2—second wash; E1—first elution fraction; E2—second elution fraction; GF—sample after gel-filtration and concentration; (C)—mass spectrometer analysis.

Proteolysis occurred at the N-terminal residues because the proteins were unable to bind to the IMAC column (Figure 1B). A similar situation was previously observed in the case of SeMV $\Delta 70\text{Pro}$ [32] and in our previous experiments [17]. However, the expressed RGMoV Pro sequence did not contain any confirmed Pro cleavage sites typically used by sobemoviruses [11]. $\Delta 50\text{Pro}$, $\Delta 50\text{Pro-E/A-VPg}$, and $\Delta 117\text{Pro}^{\text{cm}}$ were purified by gel filtration on a Superdex200 column for further structural studies and measurement of the MW of the cleavage product for localization of the cleavage site. To identify the exact proteolytic site, we performed mass spectrometry (MS) analysis of the purified $\Delta 50\text{Pro}$ C-terminal domain. MS analysis identified a 20.93 kDa protein (Figure 1C) that corresponded to the protein band visible in SDS-PAGE (Figure 1B). Based on the MS data, we concluded that the protein is 67 residues shorter, and that proteolysis could occur between residues 117 and 118, resulting in a Pro domain with a total truncation of 117 aa ($\Delta 117\text{Pro}$). The predicted MW of $\Delta 117\text{Pro}^{\text{cm}}$ was 21.6 kDa, but by MS analysis, it was 21.4 kDa (Figure 1C). This demonstrates that $\Delta 117\text{Pro}^{\text{cm}}$ lacks the first M because of the following small residue [36]. A similar pattern was observed for bacteriophage Q β CP expressed in yeasts [37] and two distinct sobemovirus CPs (CfMV and RYMV) expressed in *Escherichia coli* [38].

2.2. Identification of Serine-like 3C Protease Transmembrane Domain Cleavage Site

Residues 117 and 118 in Pro correspond to E and A, respectively; therefore, we concluded that cleavage occurred at the E¹¹⁷/A¹¹⁸ site. To confirm the presence of an

E/A site in RGMoV, we mutated E¹¹⁷ to A (E117A) for both $\Delta 50$ Pro and $\Delta 50$ Pro-E/A-VPg, leaving a 6xH-tag at the N-terminus (Figure 2, No 4 and 5).

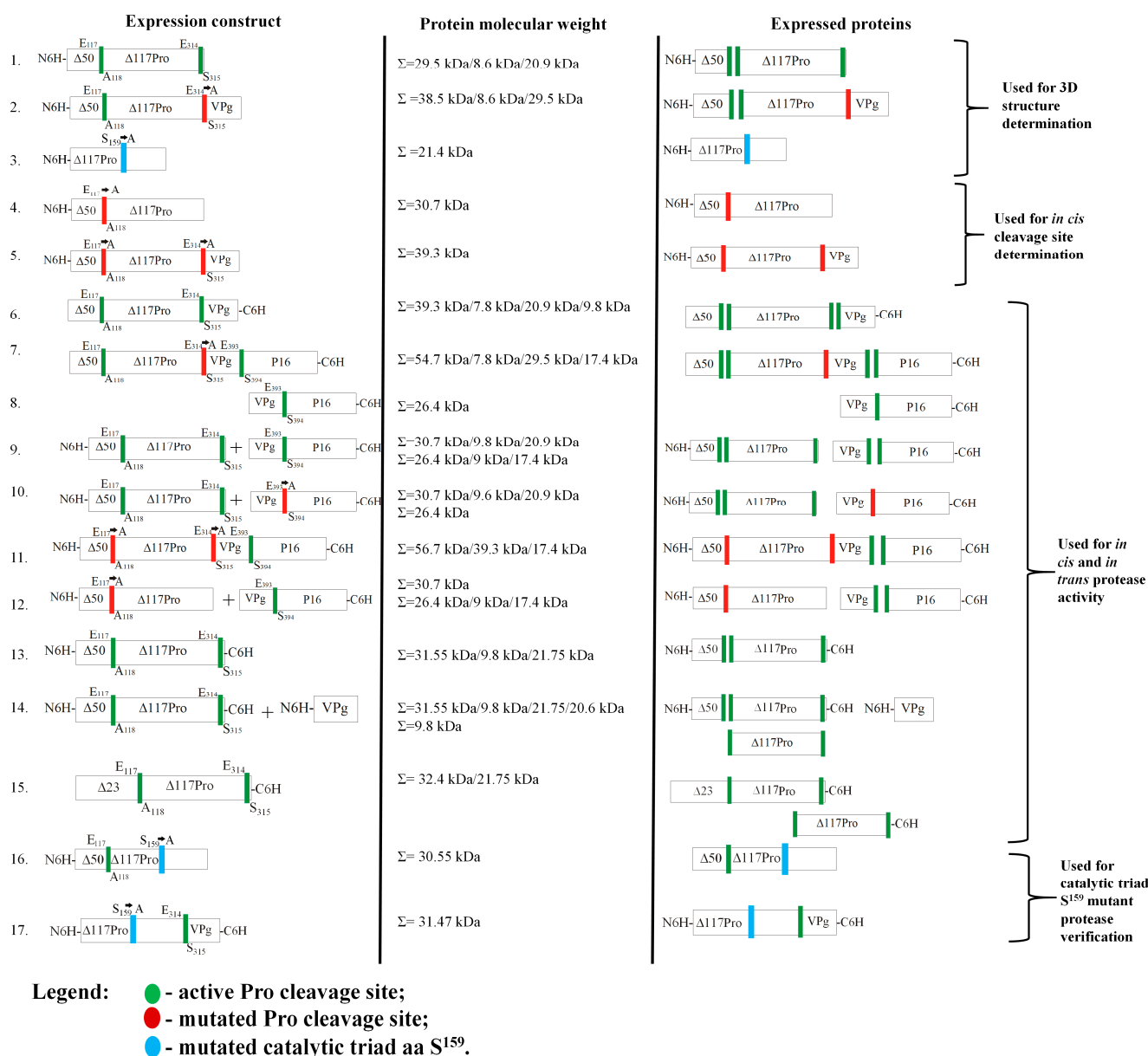


Figure 2. Schematic overview of the Pro variants used for 3D structure determination, catalytic activity experiments depending on active cleavage sites and cofactors or catalytic triad amino acid (aa) residues. 1— $\Delta 50$ Pro, active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, active cleavage site E³¹⁴/S³¹⁵ *in trans*, 6xH-tag at N-terminus; 2— $\Delta 50$ Pro-E/A-VPg, active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, inactivated cleavage site E³¹⁴/S³¹⁵ by mutation of E to A, 6xH-tag at the N-terminus of Pro; 3— $\Delta 117$ Pro^{cm}, catalytic triad S¹⁵⁹ mutated to A; 4— $\Delta 50$ Pro-TMD-N6H, inactivated cleavage site E¹¹⁷/A¹¹⁸ by mutation of E to A, active cleavage site E³¹⁴/S³¹⁵ *in trans*, 6xH-tag at N-terminus of Pro; 5— $\Delta 50$ Pro-TMD-E/A-VPg-N6H, inactivated cleavage site E¹¹⁷/A¹¹⁸ by mutation of E to A, inactivated cleavage site E³¹⁴/S³¹⁵ by mutation of E to A, 6xH-tag at N-terminus of Pro; 6— $\Delta 50$ Pro-VPg-C6H, active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, active cleavage site E³¹⁴/S³¹⁵ *in trans*, 6xH-tag at C-terminus of VPg; 7— $\Delta 50$ Pro-E/A-VPg-P16-C6H, active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, inactivated cleavage site E³¹⁴/A³¹⁵

E mutation of E to A, active cleavage site E³⁹³/S³⁹⁴ *in trans*, 6xH-tag at C-terminus of P16; 8—**VPg-P16-C6H**, active cleavage site E³⁹³/S³⁹⁴ *in trans*, 6xH-tag at C-terminus of P16; 9—**Δ50Pro-N6H** co-expression with **VPg-P16-C6H**; 10—**Δ50Pro-N6H** co-expression with **VPg-E/A-P16-C6H**, inactivated cleavage site E³⁹³/S³⁹⁴ by mutation of E to A, 6xH-tag at C-terminus of P16; 11—**Δ50Pro-TMD-E/A-VPg-P16-CN6H**, inactivated cleavage site E¹¹⁷/A¹¹⁸ by mutation of E to A, 6xH-tag at N-terminus of Pro, inactivated cleavage site E³¹⁴/S³¹⁵ by mutation of E to A, active cleavage site E³⁹³/S³⁹⁴ *in trans*, 6xH-tag at C-terminus of P16; 12—**Δ50Pro-TMD-N6H** co-expression with **VPg-P16-C6H**; 13—**Δ50Pro-NC6H**, active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, active cleavage site E³¹⁴/S³¹⁵ *in trans*, 6xH-tag at N- and C-terminus; 14—**Δ50Pro-NC6H**, active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, active cleavage site E³¹⁴/S³¹⁵ *in trans*, 6xH-tag at N- and C-terminus of Pro co-expression with **VPg-N6H**, 6xH-tag at N-terminus of VPg; 15—**Δ23Pro-C6H** active cleavage site E¹¹⁷/A¹¹⁸ *in cis*, active cleavage site E³¹⁴/S³¹⁵ *in trans*, 6xH-tag at N-terminus; 16—**Δ50Pro^{cm}-N6H**—catalytic triad S¹⁵⁹ mutated to A; 17—**Δ117Pro^{cm}-VPg-NC6H**—catalytic triad S¹⁵⁹ mutated to A, 6xH-tag at N-terminus of Pro, 6xH-tag at C-terminus of VPg. TMD: transmembrane domain; Δ50: N-terminally truncated by 50 aa; Pro: protease; VPg: virus-genome liked protein; P16: protein with 16 kDa molecular weight; red bare: represents inactivated Pro sites; green bare: represents active Pro cleavage sites; blue bar: represents catalytic triad S; box: represents protein domains.

After expression and subsequent purification by IMAC, we purified the Δ50Pro (29.5 kDa) or Δ50Pro-E/A-VPg (38.5 kDa) domains of the expected lengths, which were present in the IMAC elution fractions (Figure 3A,B). This confirmed that proteolytic cleavage was indeed prevented by mutating the E¹¹⁷/A¹¹⁸ site, and no additional cleavage site was used, even though the Pro aa sequence contains several corresponding cleavage sites [4]. In addition, we observed that the expression levels of Δ50Pro-TMD-N6H and Δ50Pro-TMD-E/A-VPg-N6H were reduced compared to those of Δ50Pro-N6H and Δ50Pro-E/A-VPg-N6H. Δ50Pro-TMD-E/A-VPg-N6H was less soluble (Figure 3A,B, line T, S, P) which could be explained by the importance of the E¹¹⁷/A¹¹⁸ site in Pro maturation.

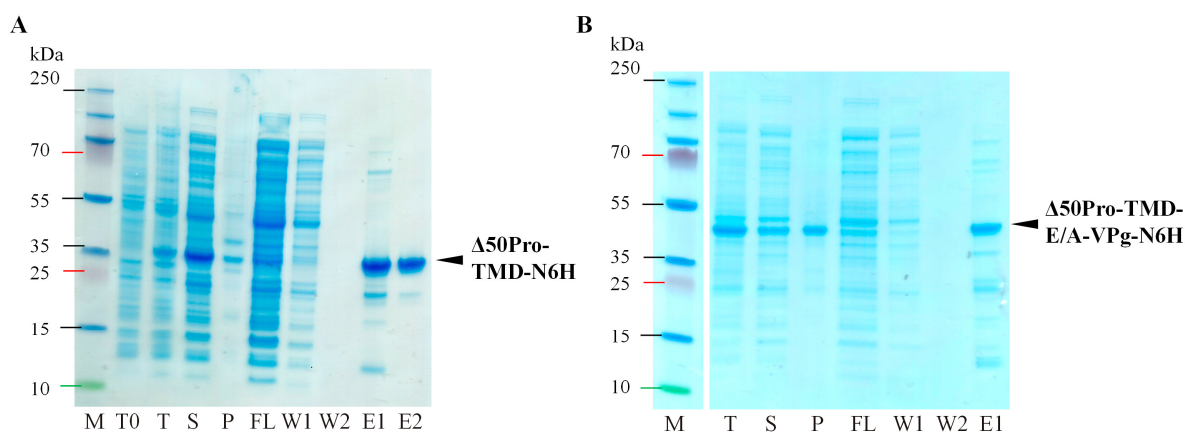


Figure 3. Δ50Pro-TMD-N6H and Δ50Pro-TMD-E/A-VPg-N6H expression purification analysis. (A)—expression and purification analysis of Δ50Pro-TMD-N6H; (B)—expression analysis of Δ50Pro-TMD-E/A-VPg-N6H; M—protein marker (Page Ruler Plus, Thermo Fisher Scientific, Waltham, MA, USA); T0—total cell lysate before expression; T—total cell lysate after expression; S—soluble protein fraction; P—pellet; FL—flow fraction after IMAC; W1—first wash; W2—second wash; E1—first elution fraction; E2—second elution fraction; arrow indicates expressed Pro.

The expression of Δ23Pro with the C-terminal 6xH-tag was low (32.4 kDa; Figure 2, No 15; Figure 4), and the protein was mostly insoluble, but proteolysis *in cis* was observed by purification of Δ117Pro-C6H in small amounts (21.75 kDa; Figure 2, No 15; Figure 4A, line E1), which was detected by Western blotting (WB) with antibodies against a 6xH-tag (Figure 4B, line E1).

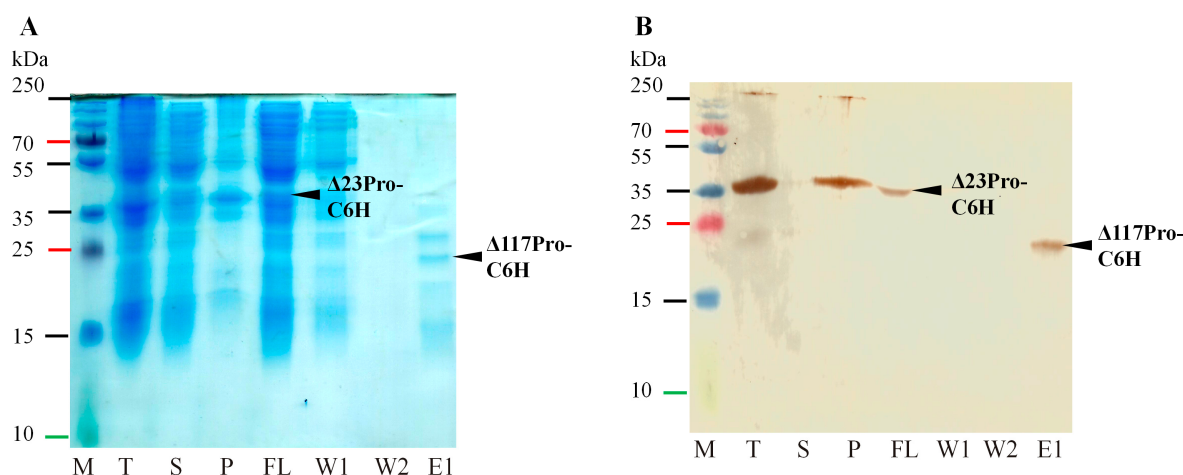


Figure 4. $\Delta 23\text{Pro-C6H}$ expression and purification analysis in SDS-PAGE and WB. (A)—expression and purification analysis of $\Delta 23\text{Pro-C6H}$ in SDS-PAGE; (B)—expression and purification analysis of $\Delta 23\text{Pro-C6H}$ in WB with 6xH-tag antibodies in ratio 1:1000; M—protein marker (Page Ruler Plus, Thermo Fisher Scientific, Waltham, MA, USA); T—total cell lysate after expression; S—soluble protein fraction; P—pellet; FL—flow fraction after IMAC; W1—first wash; W2—second wash; E1—first elution fraction.

The alignment of RGMoV Pro N-terminal part with 20 sobemoviruses Pro sequences demonstrated that the E/A cleavage site is also present in other sobemoviruses, including Rottboellia yellow mottle virus, RYMV, Imperata yellow mottle virus, Cymbidium chlorotic mosaic virus, Cymbidium chlorotic mosaic virus, and lucerne transient streak virus, which confirms the possibility of utilizing other cleavage site variants [34]. Alignment also revealed other possible cleavage site sequences. In Artemisia virus A and velvet tobacco mottle virus, we identified an E/G (glutamic acid/glycine) cleavage site. Additionally, we identified an E/C site in four cases, papaya lethal yellowing virus, Rubus chlorotic mottle virus, sowbane mosaic virus, and subterranean clover mottle virus, and an E/M site in CfMV (E/T and E/N cleavage sites is in close proximity [34]), but none of them were tested *in vivo*. In the remaining sobemoviruses, we identified previously described cleavage sites, such as E/S, in five viruses, namely SBMV, SeMV, Solanum nodiflorum mottle virus, soybean yellow common mosaic virus, and SCPMV, and E/T in turnip rosette virus and blueberry shoestring virus (Figure S1). Considering the controversial results from SeMV Pro that we obtained, we were eager not only to test Pro activity *in cis* and *trans* in different combinations with ORF2a encoded proteins, but also to confirm the proteolytic cleavage sites used for RGMoV ORF2a.

2.3. Confirming of ORF2a Proteolysis Cleavage Sites

To confirm ORF2a and to test the *E. coli* system for Pro cleavage site identification and confirmation, we created two expression vectors. One contained a native $\Delta 50\text{Pro-VPg}$ fusion with 6xH-tag at the C-terminus of VPg (Figure 2, No 6) and $\Delta 50\text{Pro-E/A-VPg}$ fused with the P16 domain (equivalent to the SeMV P18 domain) with 6xH-tag at the P16 C-terminus (Figure 2, No 7). SDS-PAGE analysis revealed the cleavage of separate VPg-C6H or P16-C6H domains, which were subsequently purified by IMAC (Figure 5A,B).

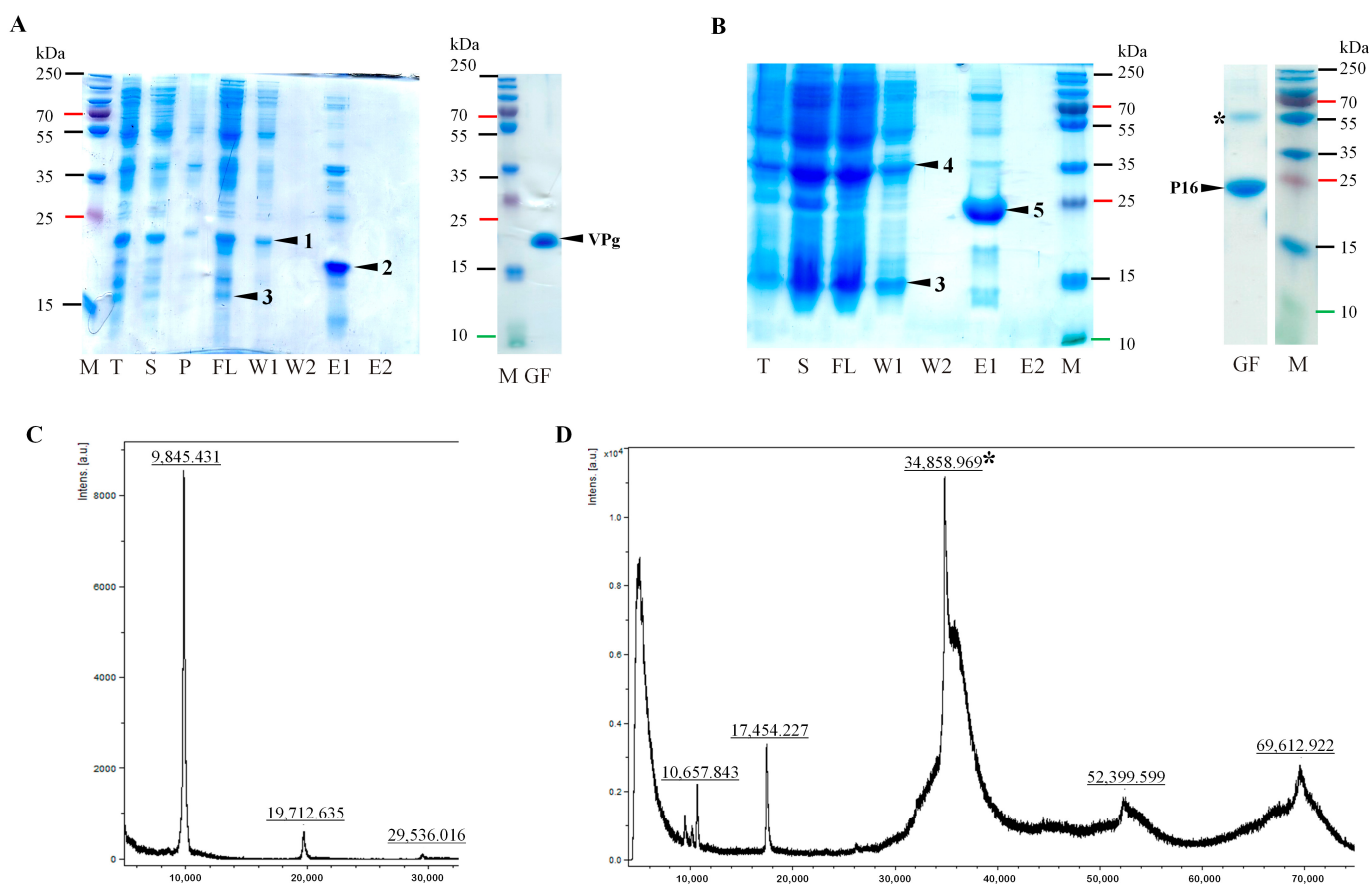


Figure 5. Expression and purification analysis of VPg-C6H and P16-C6H from polyprotein processing. (A)—expression and purification analysis of $\Delta 50\text{Pro-VPg-C6H}$ in SDS-PAGE; (B)—expression and purification analysis of $\Delta 50\text{Pro-E/A-VPg-P16-C6H}$ in SDS-PAGE; (C)—mass spectrometer analysis of VPg-C6H; (D)—mass spectrometer analysis of P16-C6H; M—protein marker (Page Ruler Plus, Thermo Fisher Scientific, Waltham, MA, USA); T—total cell lysate after expression; S—soluble protein fraction; P—pellet; FL—flow fraction after IMAC; W1—first wash; W2—second wash; E1—first elution fraction; 1—C-terminus of proteolytically truncated $\Delta 50\text{Pro}$ ($\Delta 117\text{Pro}$); 2—VPg-C6H; 3—N-terminus of proteolytically truncated $\Delta 50\text{Pro}$; 4—C-terminus of $\Delta 50\text{Pro-E/A-VPg}$ ($\Delta 117\text{Pro-E/A-VPg}$); 5—P16-C6H; GF—sample after gel filtration and concentration; *—P16-C6H dimer.

Elution fractions were additionally purified by gel filtration on a Superdex200 column for analysis with MS to confirm the MW of proteins, because VPg and P16 showed abnormal behavior in SDS-PAGE, and to identify Pro cleavage sites (Figure 5A,B). A similar situation was observed for CfMV VPg, as it was isolated from the host organism and detected in WB as a 24 kDa protein and not a 12 kDa protein, as expected [39]. The shift in mobility in SDS-PAGE could be caused by the acidic nature of VPg proteins (IP~4). This abnormal mobility has been reported previously [40]. A similar behavior of the P16 homolog has been reported earlier [17]. This unusual behavior of P16 in SDS-PAGE could be explained by its high isoelectric point (IP = 9.26). Positively charged proteins can also behave anomalously on SDS-PAGE, as in the case of the SeMV P8 protein [41]. VPg-C6H MS analysis of MW confirmed that in VPg, the E/S cleavage site was used. This has also been described for another native VPg [13]. The predicted MW of cleaved VPg-C6H MW was 9.87 kDa as revealed by MS analysis (MW 9.85 kDa; Figure 5C), which is in good agreement with the predicted MW and E/S sites. In the P16-C6H case, MS analysis confirmed the usage of the E/S site ($\text{E}^{393}/\text{S}^{394}$). To confirm this and to exclude any possibility of an additional preferable cleavage site, as was described for SeMV [41], we mutated the cleavage site E^{393} to A between the VPg-P16 domains ($\text{E}393\text{A}$; Figure 2, No 10). Co-expression with $\Delta 50\text{Pro-N6H}$ analysis by SDS-PAGE did not show any cleavage of VPg from P16 (Figure 6C), thereby

confirming the previously identified cleavage site. In addition, no additional processing products that could be confirmed were detected. Therefore, this raises the question of whether the P18 domain of SeMV undergoes processing *in cis* [41]. As demonstrated in our previous experiments, RGMoV P16 tends to oligomerize and can be classified as a natively unfolded protein [17]. This experiment also revealed that the *E. coli* expression system is acceptable for Pro cleavage site mapping and testing as a fast and inexpensive preliminary method.

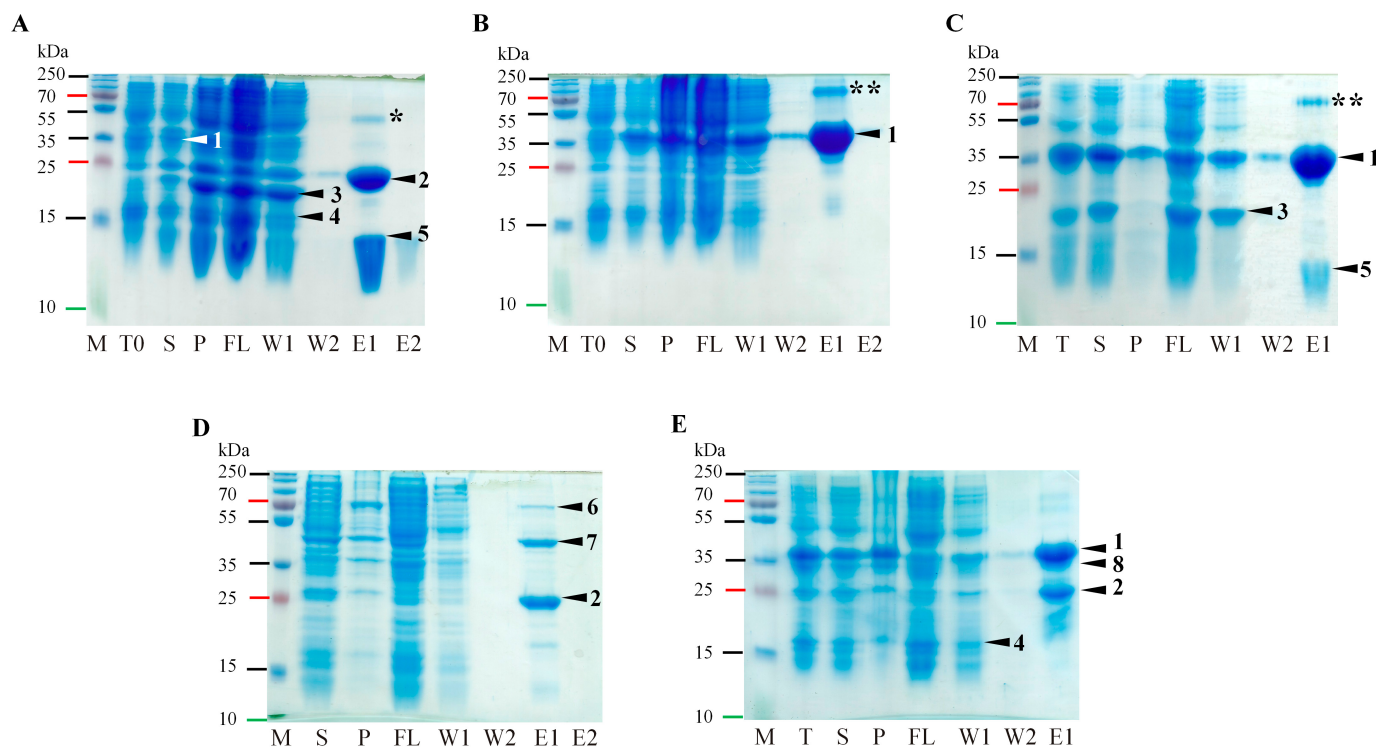


Figure 6. Pro *in trans* cleavage test and VPg-P16 cleavage site mapping. (A)—co-expression and purification analysis of $\Delta 50\text{Pro-N6H}$ and VPg-P16-C6H in SDS-PAGE; (B)—expression and purification analysis of VPg-P16-C6H in SDS-PAGE; (C)—co-expression and purification analysis of $\Delta 50\text{Pro-N6H}$ and VPg-E/A-P16-C6H in SDS-PAGE; (D)—expression and purification analysis of $\Delta 50\text{Pro-TMD-E/A-VPg-P16-NC6H}$; (E)—co-expression and purification analysis of $\Delta 50\text{Pro-TMD-E/A-N6H}$ and VPg-P16-C6H in SDS-PAGE; M—protein marker (Page Ruler Plus, Thermo Fisher Scientific, Waltham, MA, USA); T—total cell lysate after expression; S—soluble protein fraction; P—pellet; FL—flow fraction after IMAC; W1—first wash; W2—second wash; E1—first elution fraction; 1—VPg-P16-C6H; 2—P16-C6H; 3—C-terminus of proteolytically truncated $\Delta 50\text{Pro}$; 4—VPg; 5—N-terminus of proteolytically truncated $\Delta 50\text{Pro}$; 6— $\Delta 50\text{Pro-TMD-E/A-VPg-P16-NC6H}$; 7— $\Delta 50\text{Pro-TMD-E/A-VPg}$; 8— $\Delta 50\text{Pro-TMD-N6H}$; *—P16-C6H dimer, **—VPg-P16-C6H dimer.

2.4. Serine-like 3C Protease *trans*- and *cis*-Activity Tests in the Presence and Absence of VPg

We were curious if RGMoV Pro was active only *in trans* when it was fused with VPg, as described for SeMV. To test our hypothesis, we tested the *in cis* and *trans* activities of Pro in a two-plasmid system. For the *in trans*-cleavage test, we expressed $\Delta 50\text{Pro}$ as a single domain from the pCold1 plasmid, and a VPg-P16 domain with a 6xH-tag at the C-terminus of P16 from the pET28a(+) plasmid (Figure 2, No 9). Total cell lysate analysis revealed that Pro could cleave the VP-P16-C6H fusion domain (Figure 6A) into VPg and P16-C6H, showing a similar pattern in SDS-PAGE as in the case of $\Delta 50\text{Pro-E/A-VPg-P16-C6H}$, where the protein with a MW corresponding to that of P16-C6H was purified by IMAC (Figure 5B). As a control, we expressed the VPg-P16-C6H fusion protein separately (Figure 2, No 8; Figure 6B) and its E393A mutant in co-expression with $\Delta 50\text{Pro}$ (Figure 2, No 10; Figure 6C). These results indicate that RGMoV Pro does not lose its *trans* and *cis* activity when VPg is

expressed as a separate domain. This is in contrast to previous *in vitro* experiments, which showed that Pro is active only when VPg is fused to its C-terminal part [14]. This indicated that the RGMoV Pro cleavage mechanisms were different.

In the SeMV Pro E133A mutant, Nair and colleges proposed that its activity is reduced, and P18 processing into P10 and P8 is corrupted [41]. Therefore, we performed a Pro activity test by co-expressing $\Delta 50$ Pro-TMD-N6H with VPg-P16-C6H (Figure 2, No 12; Figure 6E), and created a $\Delta 50$ Pro-TMD fusion construct containing a mutated cleavage site between the Pro and VPg domains and active sites between the VPg and P16 domains and the 6xH-tag at the N- and C-termini ($\Delta 50$ Pro-TMD-E/A-VPg-P16-NC6H; Figure 2, No 11; Figure 6D). As revealed by SDS-PAGE analysis, Pro was still active *in trans*, cleaving off separate P16-C6H domains in both cases, but in a reduced manner (Figure 6D,E), highlighting the importance of the appropriate matured form of Pro during polyprotein processing in terms of speed.

In addition, in this case, no P16 processed variant was observed. In the SeMV case, it was possible to identify a weak signal of the 14 kDa protein in WB as the N-terminal part of $\Delta N70$ Pro with 6xH-tag, which could be recognized by polyclonal antibodies raised against P8 with 6xH-tag [41]. As revealed by SDS-PAGE analysis, in RGMoV, $\Delta 50$ Pro with the 6xH-tag N-terminal part after cleavage in SDS-PAGE migrates abnormally as a ~15 kDa protein instead of an 8.6 kDa protein (Figure 1B). This was due to its acidic nature (IP-4.61). The SeMV Pro N-terminal part is also acidic (IP-4.29), and its MW (7 kDa) is close to the mass of P8; therefore, it is difficult to confirm that it is indeed P8. These *in vivo* experiments revealed that the RGMoV C-terminal domain after VPg does not undergo processing into two separate domains, despite several Pro cleavage sites being used *in trans* or *cis* [4], and our previous experiments confirmed P16 as a monolithic domain [17].

To test $\Delta 50$ Pro *cis* activity in the absence of VPg, we expressed $\Delta 50$ Pro with a cleavage site before the 6xH-tag sequence at the C-end (Figure 2, No 13). SDS-PAGE analysis revealed that Pro is active *in cis*, releasing $\Delta 117$ Pro, which can be purified by IMAC (Figure 7A). In addition, $\Delta 117$ Pro released from $\Delta 50$ Pro, used in 3D determination and MS analysis, confirmed the presence of an *in trans* cleavage site with two additional aa—ESSN—in the Pro C-terminal region (Figure 1). This confirmed that Pro has *in cis* activity without VPg.

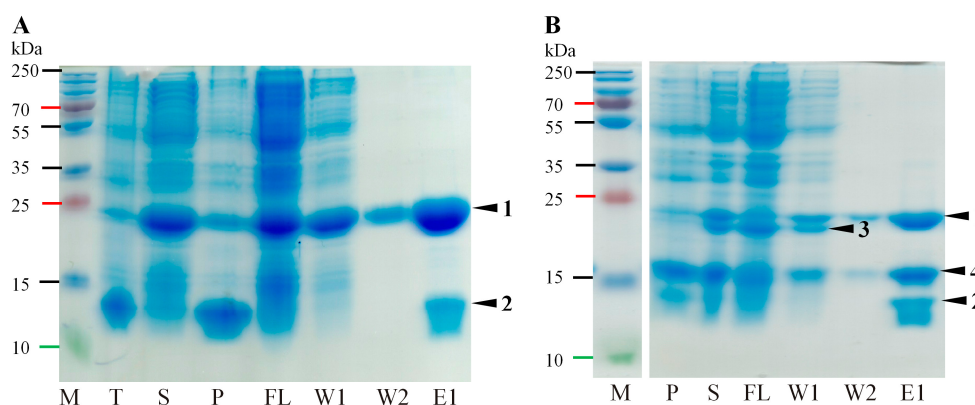


Figure 7. Pro *in cis* and *in trans* cleavage test without and with VPg-N6H. (A)— $\Delta 50$ Pro-NC6H expression and purification analysis in SDS-PAGE; (B)—co-expression and purification analysis of VPg-N6H and $\Delta 50$ Pro-NC6H in SDS-PAGE; M—protein marker (Page Ruler Plus, Thermo Fisher Scientific, Waltham, MA, USA); T—total cell lysate after expression; S—soluble protein fraction; P—pellet; FL—flow fraction after IMAC; W1—first wash; W2—second wash; E1—first elution fraction; 1— $\Delta 117$ Pro-N6H; 2—N-terminus of proteolytically truncated $\Delta 50$ Pro; 3— $\Delta 117$ Pro; 4—VPg-N6H.

This construct was also used for co-expression with VPg to test for *trans* activity. Co-expression analysis of total cell lysates revealed an additional protein band (Figure 7B). After sample analysis of the purified proteins by IMAC, the newly appeared protein band did not bind to the column material. This confirms that the *in trans* activity of Pro is mediated by the presence of VPg and the importance of protein–protein interactions.

The expression of the Pro^{cm} variant confirmed the importance of S¹⁵⁹ for the catalytic activity of RGMoV Pro (Figures S2 and S3).

2.5. Crystal Structures of $\Delta 117$ Pro, $\Delta 117$ Pro-E/A-VPg, and $\Delta 117$ Pro^{cm}

We crystallized and solved 3D crystal structures for two catalytically active forms of $\Delta 117$ Pro: one in its VPg-free form, the second in its VPg-bound state with fused VPg at the C-terminal part ($\Delta 117$ Pro-E/A-VPg), and one catalytically defective $\Delta 117$ Pro^{cm} form with S¹⁵⁹/A mutation. The overall fold change of $\Delta 117$ Pro is typical for CLP and is very similar to the structure of SeMV Pro (Figure 8).

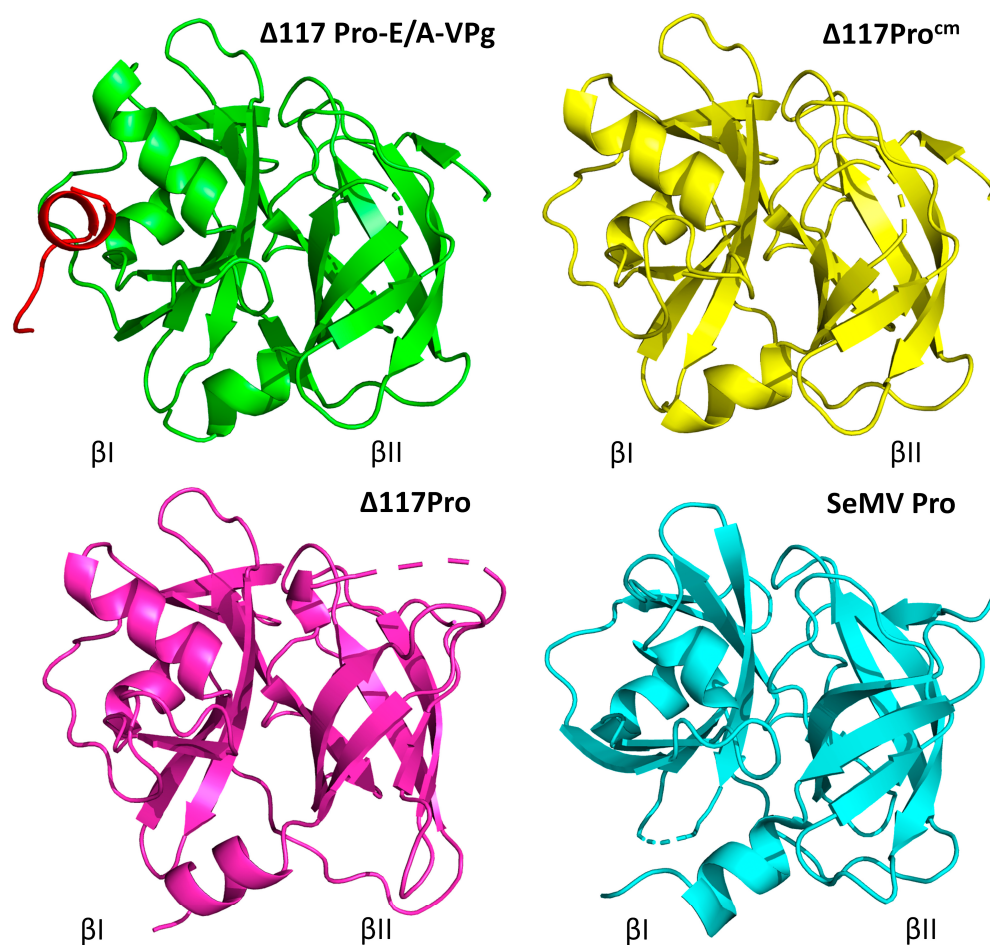


Figure 8. 3D crystal structures of $\Delta 117$ Pro-E/A-VPg (PDB ID 6FF0, green) and VPg sequence GEMTWADMVE (red), $\Delta 117$ Pro^{cm} (PDB ID 7YZV, yellow), RGMoV $\Delta 117$ Pro (PDB ID 6FEZ, magenta), and SeMV Pro domain (PDB ID 1ZYO, cyan). Molecule sides containing β I and β II barrels are marked.

The VPg sequence GEMTWADMVE, corresponding to residues 256–265 in $\Delta 117$ Pro-E/A-VPg, forms a short α -helix that binds to the surface of $\Delta 117$ Pro (Figure 9A). W²⁶⁰ is a key residue in this interaction; it is located in a very tight and largely hydrophobic pocket, forming a hydrogen bond with the main-chain carbonyl oxygen of D⁷⁷. The interaction of the VPg helix with $\Delta 117$ Pro is also ensured by two additional hydrogen bonds: one between the K⁵² side chain nitrogen and V²⁶⁴ main chain oxygen, and the second between the I⁸² main chain nitrogen and M²⁵⁸ main chain oxygen (Figure 9B). VPg-binding sites are mostly hydrophobic, as they are formed by L⁵¹, M⁵⁵, M⁷⁶, W⁸⁰, I⁸², L⁸⁷, I⁹⁴, and V⁹⁶ (Figure 9C). The VPg helix is somewhat amphipathic, having W²⁶⁰ and V²⁶⁴ side chains located towards the VPg site, and T²⁵⁹, D²⁶², and E²⁶⁵ side chains exposed to the outside environment, so the VPg–Pro interaction is also strengthened by hydrophobic interactions.

Curiously, SeMV Pro contains an analogous hydrophobic pocket, but it very likely does not have a similar function, mainly because W¹⁸³ already occupies its center (Figure 9D).

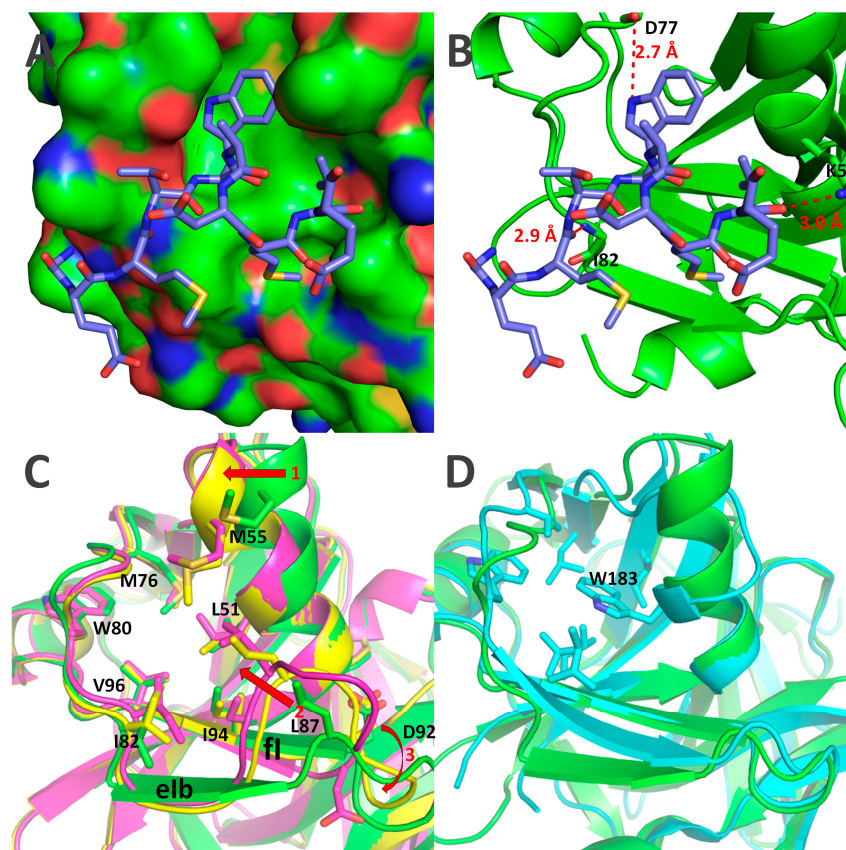


Figure 9. RGMoV VPg and its Δ117Pro binding site on βI barrel. (A)—a surface model of Δ117Pro-E/A-VPg with bounded VPg GEMTWADMVE motif, corresponding to 256–265 aa (stick model). In the surface model, the carbon atoms are displayed in green, oxygen atoms in red, and nitrogen atoms in blue. (B)—hydrogen bonds formed between VPg and Δ117Pro. Distances are measured in angstroms. (C)—aa forming RGMoV Pro VPg binding pocket and the conformational shifts in the structure of α1 helix (shift 1), in the elb-fl loop (shift 2), and the positioning of catalytic triad D⁹² (shift 3). Shifts are indicated with red arrows. (D)—comparison of RGMoV Δ117Pro-E/A-VPg (PDB ID 6FF0, green) and SeMV Pro domain (PDB ID 1ZYO, cyan). The hydrophobic pocket of the SeMV Pro domain is indicated by stick modelling.

The binding of VPg causes several significant conformational shifts in the Δ117Pro structure (Figures 4C and 5).

First, in the absence of VPg, L⁸⁷ shifted towards the VPg-binding pocket. Consequently, the elb β-strand is also pulled towards the VPg binding site, disrupting its β-strand conformation and significantly altering the position of the loop between the elb and fl strands. Second, the α1 helix in the VPg-free structure is also slightly shifted towards the empty VPg binding site (Figure 9C); although, the position of the catalytic triad H⁴⁹, located on this helix, is minimally affected. These two conformational shifts are driven by the presence or absence of VPg, since they align well in the Δ117Pro^{cm} and Δ117Pro cases, which do not contain VPg at all.

There were several other conformational shifts for which no clear causation between the presence and absence of VPg could be ascertained. The rotation of the catalytic triad D⁹² (Figure 9C) is one such feature. In Δ117Pro chain A, it is rotated away from the catalytic triad H⁴⁹, thereby disrupting the functionality of the catalytic triad since D⁹² is too far from H⁴⁹ to ensure any proton transfer. However, in the Δ117Pro chain B, Δ117Pro^{cm}, and Δ117Pro-E/A-VPg cases, D⁹² is directed towards H⁴⁹, as would be expected in an active

Pro. There are several more conformational shifts unique only to $\Delta 117$ Pro (Figure 10) and that are not present in $\Delta 117$ Pro^{cm} and $\Delta 117$ Pro-E/A-VPg. These are the partial unwinding of the β II barrel and subsequent movement of the bII and cII β -strands and the cII/dII loop, and the disordering of the eII/fII loop (Figure 10A). These conformational shifts may affect the positioning of G¹⁵⁷ (Figure 10B). This residue, judging by comparison with the 3D structures of other serine Pros, is responsible for the formation of the oxyanion hole, and its disruption in the $\Delta 117$ Pro structure could directly modulate the catalytic activity. These conformational shifts also had a significant influence on the position of conserved key residues forming the Glu-binding site (H¹⁵⁰, H¹⁷³, T¹⁵⁴, and N¹⁸³; Figure 10C,D). While H¹⁷³ and N¹⁸³ did not change their positions significantly, H¹⁵⁰ and T¹⁵⁴ were noticeably shifted due to the movement of bII-cII strands. This shift could also render the Glu-binding site unsuitable for cleavage site recognition.

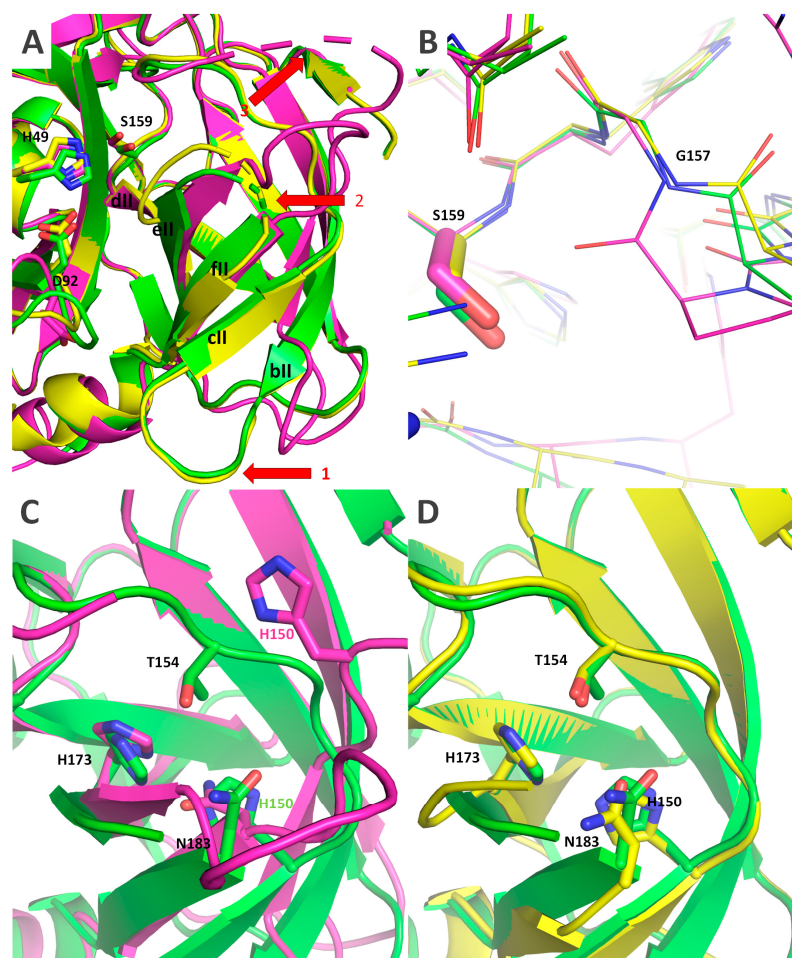


Figure 10. The conformational transition between RGMoV $\Delta 117$ Pro VPg-bound and VPg-free states in β II barrel. (A)—conformational shifts of bII and cII β -strands (shift 1), eII/fII loop (2), and cII/dII loop (3) in $\Delta 117$ Pro-E/A-VPg (PDB ID 6FF0; green), $\Delta 117$ Pro (PDB ID 6FEZ; magenta), and $\Delta 117$ Pro^{cm} (PDB ID 7YZV; yellow) structures. (B)—possible disruption of oxyanion hole formed by G¹⁵⁷ in $\Delta 117$ Pro (PDB ID 6FEZ; magenta) compared to $\Delta 117$ Pro-E/A-VPg (PDB ID 6FF0; green) and $\Delta 117$ Pro^{cm} (PDB ID 7YZV; yellow) structures. (C)—influence of the conformational shift in eII/fII and cII/dII loops on the positions conservative key residues in Glu-binding site (H¹⁵⁰, H¹⁷³, T¹⁵⁴ and N¹⁸³) in $\Delta 117$ Pro (PDB ID 6FEZ; magenta) and $\Delta 117$ Pro-E/A-VPg (PDB ID 6FF0; green) structures. (D)—influence of the conformational shift in eII/fII and cII/dII loops on the positions conservative key residues in Glu-binding site (H¹⁵⁰, H¹⁷³, T¹⁵⁴ and N¹⁸³) in $\Delta 117$ Pro^{cm} (PDB ID 7YZV; yellow) and $\Delta 117$ Pro-E/A-VPg (PDB ID 6FF0; green) structures.

None of these conformational shifts were observable in SeMV Pro; both the catalytic triad and Glu-binding regions aligned almost perfectly with $\Delta 117$ Pro-E/A-VPg (Figure 11). Such differences can be partially explained by the C²⁷⁷-C²⁴⁸ disulfide bond in the SeMV Pro structure (Figure 11). This disulfide anchors the cII strand and does not allow for movement as in the $\Delta 117$ Pro case. Curiously, RGMoV contains one C in this position, C¹⁵², but the other corresponding residue is A¹²⁴, so there is no comparable stabilizing influence.

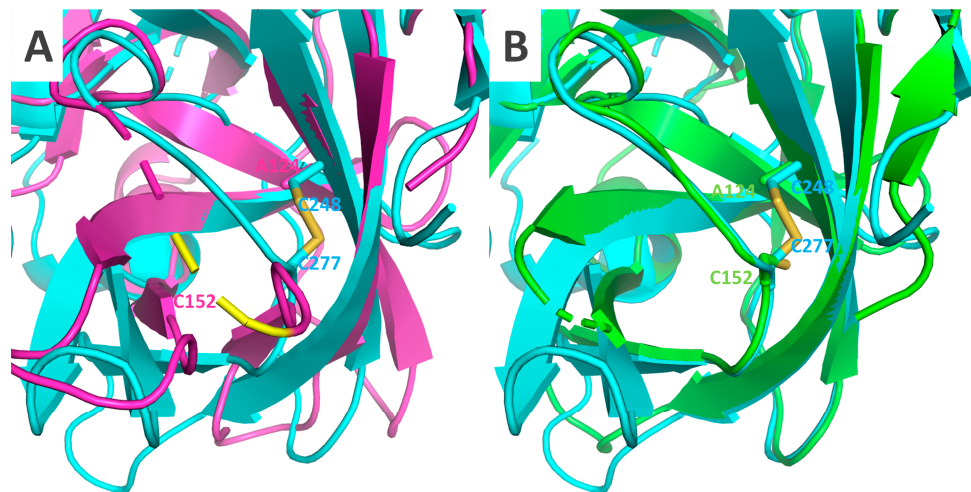


Figure 11. SeMV Pro domain (PDB ID 1ZY0, cyan) substrate binding site and C²⁷⁷-C²⁴⁸ disulfide bridge. SeMV Pro domain is aligned with $\Delta 117$ (magenta, (A)) and $\Delta 117$ -Pro-E/A-VPg (green, (B)).

Nevertheless, the C²⁷⁷-C²⁴⁸ disulfide bridge was demonstrated for SeMV Pro as non-essential for catalytic activity; therefore, its absence could be relatively unimportant [12,32]. For RGMoV Pro, there is no intermolecular disulfide bond present in SeMV Pro, and, as a result, the former is a monomer. This former SeMV intramolecular disulfide could also be unimportant because it has been speculated to be an artifact of crystallization [32]. However, RGMoV $\Delta 117$ Pro has an internal C³⁸-C¹⁹⁵ disulfide bridge that is not present in SeMV Pro. This illustrates that the Pro domain disulfide bridges, although potentially important for conformations, are not particularly conserved among different viruses. Potential disruption of the RGMoV Pro internal C³⁸-C¹⁹⁵ disulfide bridge by 20 mM dithiothreitol (DTT) led to the inactivation of proteolytic activity (Figure S4g), so the importance of this particular bridge in the activity of RGMoV Pro remains open.

3. Discussion

Sobemoviral Pros are known to cleave E/T, E/S, and E/N residues [39,41,42], but all cleavage sites are characterized only for SeMV [41]. Only cleavage sites between Pro and VPg were identified for CfMV, SBMV, RYMV, and RGMoV [8,13,39]. Sequence comparison of TMD Pro cleavage sites revealed the presence of an E/A cleavage site, which was confirmed by site-directed mutagenesis in RGMoV. Six more sobemoviruses were identified that could use the same cleavage site and probably also the E/C, E/G, and E/M sites (Figure S1). The VPg-P16 domain in RGMoV was processed using an E/S cleavage site identical to that of the Pro-VPg site. However, we did not observe further processing of the P16 domain into two subdomains, as described previously for SeMV [41], and we propose that it is not processed in *E. coli* in the SeMV case. The identified protein in WB could be the $\Delta 70$ NPro N-terminal domain containing 6xH-tag as P8, against which polyclonal antibodies were raised. Our experiments revealed that a bacterial expression system combined with MS analysis and site-directed mutagenesis is suitable for the identification of Pro cleavage sites. In addition, this system allows the testing of both *in cis* and *trans* Pro activities. Pro activity *in trans* can be achieved with free VPg co-expression without fusion, as described previously [14]. Unfortunately, we were unable to develop a FRET system for Pro catalytic

activity *in cis* to test and evaluate VPg's impact on the catalytic speed. However, we were able to detect peptide cleavage by robust MS analysis (Figure S4). Additionally, we performed an *in vitro* cleavage test using $\Delta 50\text{Pro}^{\text{cm}}$ as the substrate and added $\Delta 117\text{Pro}$ with or without VPg to monitor the possible VPg in the *cis* cleavage speed modulation of Pro. We observed that VPg did not increase the Pro cleavage speed. We also observed that Ca^{2+} and Zn^{2+} acted as inhibitors, whereas Mg^{2+} only slightly reduced Pro activity (Figure S5). Calcium is an essential nutrient involved in growth and development and is an important secondary messenger molecule. The Ca^{2+} under normal conditions is in the range of 10–200 nM; whereas the concentration of Ca^{2+} in the cell wall, vacuole, endoplasmic reticulum, and mitochondria is 1–10 mM, during stress, it can reach the micromolar level [43]. As viral infection can trigger plant stress signaling pathways, an increase in Ca^{2+} levels could be an indirect defense mechanism against viral infection. We hypothesize that this mechanism could have a favorable effect on viral particle maturation, as sobemovirus CP binds Ca^{2+} for particle stabilization [44]. In addition, Zn is involved in plant responses to pests and diseases [45]. As sobemovirus P1 is classified as a Zn-finger protein [46], it acts as a suppressor and activator of gene silencing [47] and is an infectivity factor involved in both viral replication and spreading [48], where the P1 Zn binding site may function as a redox sensor of the plant redox status to adapt its functions [49]. An increase in both elements in plant cells likely deactivates Pro, restricting polyprotein processing and reducing viral gRNA and sgRNA replication. However, the toxic effects of Ca^{2+} and Zn^{2+} can lead to cell death, thereby favoring systemic viral infection. Nevertheless, our data demonstrate that the presence or absence of RGMoV VPg does not have such a dramatic influence on *cis* proteolytic activity, as was the case with SeMV. We cannot exclude the modulating influence of RGMoV VPg on Pro activity *in planta*. However, complete loss of *cis* activity was not observed.

It has been noted previously that VPg presence acts as a switch between *in cis* and *in trans* active states [14,32]. Structural changes in sobemovirus Pro during interaction with VPg have been previously analyzed only with the CD spectrum by mutating key residues [15]. A more recent NMR study demonstrated that the interaction of VPg and Pro domains in SeMV is driven by a patch of hydrophobic residues on the surface of VPg and Pro domains [20]. Crystallography allowed us to identify visual changes in Pro structure during the interaction with VPg. We determined that a 10-residue-long VPg motif could assume the structure of an α -helix and bind to a specific site in the Pro domain. This interaction triggers major structural changes in the catalytic triad and substrate binding sites. These changes could certainly be the underlying mechanism for activation or at least stabilization of the *trans* activity of RGMoV Pro, but this mechanism is not universal among all sobemoviruses. Similar major structural changes have not been observed for the closely related SeMV Pro, the *in cis* active structure, which is very similar to VPg bound, a completely *in trans*-active RGMoV Pro. Although the interaction between Pro and VPg is driven by hydrophobic interactions in both RGMoV and SeMV, the responsive VPg structures— α - β - β - β domain for SeMV Pro and single α helix for RGMoV—are completely different. The RGMoV VPg motif WAD, with the key residue W^{260} according to PsiPred [50] and AlphaFold [51], is located in a helix, but in the SeMV case, the corresponding residues are located in a coil (Figure S6). The previously described essential W^{43} for SeMV VPg function [14,15] is located in different structural element strands, and there is no analog W residue for RGMoV in the corresponding region. Nevertheless, it can be concluded that both RGMoV and SeMV VPg structures contain an α - β - β - β domain in their N-terminal region, but their C-terminal parts after this domain are much more divergent, with RGMoV VPg containing several additional predicted α -helices and SeMV VPg being much more disordered. Overall, our research has identified novel interacting surfaces and key residues for VPg interaction with Pro, and demonstrated that the nature of this interaction and its implications are not conservative and could be highly different among sobemoviruses. The true meaning of the RGMoV Pro interaction with VPg remains controversial, since it is not

directly involved in controlling the catalytic activity of Pro *in cis*, but is required *in trans*. Further studies are required to confirm this hypothesis.

Surprisingly, conformational differences were observed between the $\Delta 117$ Pro, $\Delta 117$ Pro-E/A-VPg, and $\Delta 117$ Pro^{cm} states, with the only difference being the S¹⁵⁹/A mutation. Different crystallization conditions could be at play, but it is very likely that these differences are actual representations of natural conformational oscillations in the Pro molecule.

4. Materials and Methods

4.1. Gene Amplification and Cloning

The cDNA clone of RGMoV MAFF. No. 307043 from Japan (GenBank: EF091714; [4]) was used as the source of the DNA matrix. As sobemovirus Pro at its *N*-terminal part contains a TMD sequence, for soluble RGMoV Pro expression in *E. coli* [12], we created an *N*-terminally truncated Pro, which lacks the first 50 aa ($\Delta 50$ Pro) and contains a 6xH-tag at the *N*-terminus originating from the chosen expression vector [17]. The PCR product of $\Delta 50$ Pro was first cloned into the cloning vector pTZ-57 using the InsTAclone PCR Cloning Kit (Thermo Fisher Scientific, Waltham, MA, USA). Clone selection with the corresponding insert was performed by restriction analysis and verified by Sanger sequencing with a BigDye Cycle Sequencing Kit and ABI Prism 3100xl Genetic Analyzer (Applied Biosystems, Waltham, MA, USA). This technique was used for all PCR products. $\Delta 50$ Pro in the pETDuet1 (Novagen, Madison, WI, USA) expression vector was cloned through *Nco*I and *Hind*III cloning sites to create the pETDu- $\Delta 50$ Pro plasmid. Clone selection was performed using a restriction analysis. The $\Delta 50$ Pro-E/A-VPg fusion cleavage site mutant ($\Delta 50$ Pro-E/A-VPg) was created as previously described for SeMV with site direct mutagenesis [14]: E³¹⁵ and E³¹⁸ (aa counting from RGMoV ORF2a first M) were mutated to A using oligonucleotides RG-SP-VPg-E/A-F and RG-SP-VPg-E/A-R, and RG-SerP-RAW-NdeI-F and RG2-VPg-HindIII-R (Table S1) to introduce cloning sites. Both possible cleavage sites were mutated to prevent unexpected VPg cleavage from Pro. The cloning strategy was as in the $\Delta 50$ Pro case only, as the expression vector was pET28a(+) (Novagen, Madison, WI, USA) using the *Nde*I and *Hind*III cloning sites, where the 6xH-tag at the *N*-terminus of Pro originated from a plasmid (pET- $\Delta 50$ Pro-E/A-VPg).

The $\Delta 50$ Pro-E/A-VPg and P16 fusion vectors with a 6xH-tag at the C-terminus ($\Delta 50$ Pro-E/A-VPg-P16-C6H) were created by cloning the VPg-P16 domain with 6xH-tag at the C-terminus (VPg-P16-C6H) PCR product which was created by introducing the oligonucleotides RG2-VPg-NcoI-F and RG-P16-C6H-HindIII-R (Table S1) into pET- $\Delta 50$ Pro-E/A-VPg through the *Eco*RI and *Hind*III restriction sites (pET- $\Delta 50$ Pro-E/A-VPg-P16-C6H) or into pET28a(+) using the *Nco*I and *Hind*III cloning sites for VPg-P16 fusion expression (pET-VPg-P16-C6H). The $\Delta 50$ Pro-VPg native fusion with a 6xH-tag at the VPg C-terminus ($\Delta 50$ Pro-VPg-C6H) was amplified using the oligonucleotides RG-SerP-RAW-NcoI-F and VPg-C-6-His-HindIII-R (Table S1) and cloned into pET28a(+) using *Nco*I and *Hind*III cloning sites ($\Delta 50$ Pro-VPg-C6H). The $\Delta 50$ Pro TMD Pro cleavage site mutant was created by site-directed mutagenesis, where E¹¹⁷ was changed to A ($\Delta 50$ Pro-TMD-E/A) using overlapping oligonucleotides RGSP-E/A-F and RGSP-E/A-R (Table S1). An expression vector, pCold1- $\Delta 50$ Pro (constructed earlier [17]), was introduced using the *Nde*I and *Xho*I restriction sites (pCold1- $\Delta 50$ Pro-TMD-E/A). For the TMD $\Delta 50$ Pro-E/A-VPg mutant ($\Delta 50$ Pro-TMD-E/A-VPg), the pColdI- $\Delta 50$ Pro-TMD-E/A vector was used, and *Cfr*42I and *Hind*III restriction sites were created after ligation of the pColdI- $\Delta 50$ Pro-TMD-E/A-VPg plasmid. In the VPg-P16-C6H cleavage site mutant, E³⁹³ was mutated to A (VPg-E/A-P16-C6H) by overlapping oligonucleotides RG-VPg-E/A-P16-F and RG-VPg-E/A-P16-R (Table S1) in the first PCR reaction and the second RG2-VPg-NcoI-F and RG-P16-C6H-HindIII-R (Table S1), and the mutant was cloned into pET28a(+) using *Nco*I and *Hind*III cloning sites, which resulted in the pET-VPg-E/A-P16 expression vector. For $\Delta 23$ Pro-C6H Pro-C-6-His-HindIII-R and RG-SerP-E/S-Nco-F, oligonucleotides (Table S1) were used. $\Delta 23$ Pro-C6H was cloned into the pTZ57 vector for sequence verification, and the correct $\Delta 23$ Pro-C6H coding sequence was cloned into pET28a(+) using *Nco*I and *Hind*III cloning sites to obtain plasmid pET-

$\Delta 23\text{Pro-C6H}$. The $\Delta 50\text{Pro-NC6H}$ construct was created based on pCold1- $\Delta 50\text{Pro}$, where a *6xH-tag* sequence was added from pET- $\Delta 23\text{Pro-C6H}$ in cloning sites *XhoI* and *HindIII*, creating the construct pCold1- $\Delta 50\text{Pro-NC6H}$.

The catalytic center $\Delta 117\text{Pro}$ mutant ($\Delta 117\text{Pro}^{\text{cm}}$) was created using oligonucleotides SP-S/A-F and SP-S/A-R (Table S1), where S¹⁵⁹ was substituted with A by site direct mutagenesis, and using SP-NCE-*HindIII*-R and SP-AAV-*NcoI*-N6H-F (Table S1), cloning sites *NcoI* and *HindIII* were introduced for cloning into the pET28a(+) expression vector (pET- $\Delta 117\text{Pro}^{\text{cm}}$ -N6H).

4.2. Protein Expression and Purification

The C2566 (New England Biolabs, Ipswich, MA, USA) or BL21(DE3) (New England Biolabs, Ipswich, MA, USA) cell strains were used for all protein expression and co-expression assays. Overnight cultures grown at 37 °C were transferred to 200 mL 2xTY media at a ratio of 1:10. The cells were cultivated at 30 °C and 200 rpm until the optical density at 600 nm reached 0.8. Cells were induced with a final concentration of 0.2 mM IPTG and cultivated at 20 °C at 200 rpm for 16 h. After cultivation, the cells were collected by slow-speed centrifugation and frozen at −70 °C.

Cells were disrupted by ultrasound using an ultrasound disintegrator UP200S (Hielscher Ultrasonics, Teltow, Germany) at period 0.5 and intensity 70% for 16 min in disruption buffer ($\Delta 50\text{Pro}$, $\Delta 50\text{Pro-E/A-VPg}$, $\Delta 117\text{Pro}^{\text{cm}}$, $\Delta 50\text{Pro-NC6H}$, $\Delta 50\text{Pro-VPg-C6H}$, $\Delta 50\text{Pro-E/A-VPg-P16-C6H}$, $\Delta 50\text{Pro-TMD-E/A}$, $\Delta 50\text{Pro-TMD-E/A-VPg}$, $\Delta 23\text{Pro-C6H}$, VPg-P16-C6H, VPg-E/A-P16-C6H, and all co-expressions 1xPBS; $\Delta 50\text{Pro}^{\text{cm}}$ 0.5 M UREA, 1xPBS, 1 M NaCl). $\Delta 50\text{Pro}^{\text{cm}}$ was purified on Protino Ni-IDA 2000 Packed Columns (Macherey-Nagel, Düren, Germany), and elution fractions containing $\Delta 50\text{Pro}^{\text{cm}}$ were dialyzed in a 6–8 kDa dialysis membrane (Spectrum Chemical, New Brunswick, NJ, USA) against 1:100 volumes of 0.5 M UREA, 1xPBS, 1 M NaCl, and 5 mM imidazole. Purification of $\Delta 50\text{Pro}$, $\Delta 50\text{Pro-E/A-VPg}$, $\Delta 117\text{Pro}^{\text{cm}}$, VPg-C6H, and P16-C6H was performed in two steps. First, we used PrepEase Histidine-tagged Protein Purification Midi Kit-High Yield (Affymetrix, Santa Clara, CA, USA) or Protino Ni-IDA 2000 packed columns (Macherey-Nagel, Düren, Germany) for $\Delta 50\text{Pro}$ and $\Delta 50\text{Pro-E/A-VPg}$ to facilitate the removal of the proteolytically cleaved N-terminal region of Pro. Gel filtration was then performed during the final cleaning step. The gel filtration process was performed on an Äkta Pure 25 XK 16/70 column packed with 120 mL Superdex 200 (GE Healthcare, Chicago, IL, USA), and purified proteins were eluted with 1xPBS at 1 mL/min by collecting 2 mL per fraction. Fractions containing purified protein were united and concentrated on Amicon-Ultra-15 10 K filtration units (Merck-Millipore, Darmstadt, Germany), and the Amicon-Ultra-0.5 10 K filtration units (Merck-Millipore, Darmstadt, Germany) were used for buffer exchange from 1xPBS to 20 mM TRIS-HCl (pH 8.0), suitable for protein crystallization. Protein concentrations were measured using a Qubit 2.0 Fluorometer with a Qubit Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA), according to the manufacturer's protocol.

Cells from co-expression experiments were treated as previously described. Soluble proteins were purified with the PrepEase Histidine-tagged Protein Purification Midi Kit-High Yield (Affymetrix, Santa Clara, CA, USA) or Protino Ni-IDA 2000 packed columns (Macherey-Nagel, Düren, Germany). Protein concentrations were measured on a Qubit 2.0 (Thermo Fisher Scientific, Waltham, MA, USA) with a Qubit Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA), according to the provided protocol.

4.3. Protein Analysis on SDS-PAGE and Western Blotting (WB)

Samples were analyzed on 12.5% SDS-PAGE or NuPAGE™ 4–12%, Bis-Tris, 1.0 mm mini protein gels (Thermo Fisher Scientific, Waltham, MA, USA). The samples for gel analysis were diluted with a 1xLaemmly sample buffer (50 mM Tris pH 7.0, 1% SDS, 50% glycerol, 5% β -mercaptoethanol (β -ME), 0.004% bromophenol blue) and heated for 10 min at 95 °C. Samples on SDS-PAGE were visualized by Coomassie blue R-250 [10% ethanol, 10% glacial acid, 0.1% R-250 (Sigma-Aldrich, St. Louis, MO, USA)] or G-250 [20% ethanol,

2% trichloroacetic acid, 0.5 g/L Coomassie G-250 (Sigma-Aldrich, St. Louis, MO, USA)] staining.

WB was performed on a semi-dry Western blot device transferred to an Amersham Protran 0.45 mm nitrocellulose (NC) blotting membrane (GE Healthcare, Chicago, IL, USA) at 52 mA for 45 min. For a protein with 6xH-tag detection on the NC membrane, the H-Tag Antibody HRP Conjugate Kit (Merck-Millipore, Darmstadt, Germany) was used according to the manufacturer's protocol.

4.4. Protein Analysis by Mass Spectrometry

Samples for mass spectrometry analysis were prepared as follows: 2 μ L of purified protein (1 mg/mL) was mixed with 2 μ L of 2% trifluoroacetic acid and 2 μ L 2,5-dihydroxyacetophenone (2,5-DHAP; Bruker Daltonics, Leipzig, Germany) matrix solution (50 μ mol 2,5-DHAP dissolved in 96% ethanol and 10 μ mol aqueous diammonium hydrogen citrate), and 1 μ L was spotted onto an MTP Anchor Chip 400/384TF (Bruker Daltonics, Leipzig, Germany) and left to crystallize. The sample analysis was performed on an "Autoflex" MALDI-TOF MS (Bruker Daltonics, Leipzig, Germany). A mix of mass calibration standards I and II or peptide molecular mass calibration standard II (Bruker Daltonics, Leipzig, Germany) was used for mass calibration.

4.5. In Vitro $\Delta 50\text{Pro}^{\text{cm}}$, $\Delta 117\text{Pro}$ and VPg Cleavage Test

For the in vitro cleavage experiments, all proteins were used in equal micromolar amounts (1:1:1). The micromolar concentrations were calculated according to $\Delta 50\text{Pro}^{\text{cm}}$ (13.7 μ M) in a 100 μ L reaction volume. These results were confirmed by a time-course experiment in which samples were withdrawn from the reaction mixture at different time points (10 min, 1 h, 2 h, 3 h, and 10 h) at 37 °C, and immediately stopped with 4xLaemmli sample buffer (50 mM Tris pH 7.0, 4% SDS, 50% glycerol, 10% β -ME, 0.004% bromophenol blue) followed by sample heating at 95 °C for 10 min. Samples were analyzed in 15 well 1.0 mm Bolt 4–12% Bis-Tris Plus mini protein gels (Thermo Fisher Scientific, Waltham, MA, USA) and stained with Coomassie blue G-250.

4.6. Crystallization, Data Collection, and Structure Determination

$\Delta 117\text{Pro}$, $\Delta 117\text{Pro-E/A-VPg}$, and $\Delta 117\text{Pro}^{\text{cm}}$ were transferred to 20 mM Tris-HCl (pH 8.0) and concentrated to 9.38 mg/mL, 2.9 mg/mL, and 18.1 mg/mL concentrations, respectively. All Pro variants were crystallized using the sitting drop method in 0.4 + 0.4 μ L or 1 + 1 μ L setups. The initial screening was performed using JCSG+, Structure screens 1 and 2, and Pact Premier screens (Molecular Dimensions, Rotherham, UK). $\Delta 117\text{Pro}$ crystals grew within 2–3 weeks, and the final crystallization conditions were 0.1 M HEPES (pH 7.5), 22.5% *m/v* PEG3350, 0.22 M LiSO_4 and 3% DMSO. $\Delta 117\text{Pro}^{\text{cm}}$ crystals grew within one week, and the final crystallization conditions were 0.18 M ammonium sulphate, 22% PEG8000, and 0.1 M sodium cacodylate (pH 6.5). $\Delta 117\text{Pro-E/A-VPg}$ crystals grew within 3–4 months in 0.1 M MIB, pH 9.0, and 25% PEG1500 conditions. Crystals were soaked in a cryoprotectant (30% glycerol in mother liquid) and frozen in liquid nitrogen. Diffraction data were collected at the MX 14.1 beamline in the BESSY II synchrotron electron storage ring operated by the Helmholtz-Zentrum Berlin (Berlin, Germany, [52]). $\Delta 117\text{Pro}$ crystals were diffracted to a 2.3 Å resolution, $\Delta 117\text{Pro}^{\text{cm}}$ crystals to a 1.6 Å resolution, and $\Delta 117\text{Pro-E/A-VPg}$ to a 2.1 Å resolution.

The data were indexed with MOSFLM [53] and scaled using SCALA [54] from the CCP4 suite [55]. The $\Delta 117\text{Pro}$ crystals had space group P2 with two chains in the asymmetric unit. The $\Delta 117\text{Pro-E/A-VPg}$ crystals had a space group $\text{P2}_1\text{2}_1\text{2}_1$ with one chain in the asymmetric unit. The $\Delta 117\text{Pro}^{\text{cm}}$ crystals had a space group C2 with one chain in the asymmetric unit.

The $\Delta 117\text{Pro-E/A-VPg}$ structure was determined by molecular replacement using MOLREP [56], using a $\Delta 117\text{Pro}$ homology model built with SWISS-MODEL [57] from the Pro domain of SeMV (PDB ID 1ZYO) [32]. The VPg GEMTWADMVE helix was built

by identifying the matching tryptophan residue in the VPg sequence, and built further by adding individual residues. The $\Delta 117\text{Pro}$ and $\Delta 117\text{Pro}^{\text{cm}}$ structures were determined in MOLREP using coordinates from the determined $\Delta 117\text{Pro-E/A-VPg}$ structure. The models were further built manually with Coot [58], and refined with REFMAC5 [59] and phenix.refine [60]. Data processing, refinement, and validation statistics are shown in Table S2. All structures were deposited in the worldwide Protein Data Bank under accession numbers 6FEZ ($\Delta 117\text{Pro}$), 6FF0 ($\Delta 117\text{Pro-E/A-VPg}$), and 7YZV ($\Delta 117\text{Pro}^{\text{cm}}$).

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ijms24065347/s1>.

Author Contributions: Conceptualization, G.K., K.T., A.Z. and I.B.; methodology, G.K. and I.B.; validation, G.K. and I.B.; formal analysis, R.L., I.K., G.R., V.Z., J.B. and I.B.; investigation, G.K., R.L., I.K., G.R., V.Z., J.B. and I.B.; resources, K.T., A.Z. and I.B.; data curation, G.K. and I.B.; writing—original draft preparation, G.K., R.L. and I.B.; writing—review and editing, G.K., K.T., A.Z. and I.B.; visualization, G.K. and I.B.; supervision, G.K., K.T., A.Z. and I.B.; project administration, I.B.; funding acquisition, I.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the European Regional Development Fund Project No.: 1.1.1.2/VIAA/3/19/462 “Structural and functional studies of Ryegrass mottle virus-encoded proteins”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The structures of RGMoV Pro variants were deposited in worldwide Protein Data Bank under accession numbers 6FEZ ($\Delta 117\text{Pro}$), 6FF0 ($\Delta 117\text{Pro-E/A-VPg}$), and 7YZV ($\Delta 117\text{Pro}^{\text{cm}}$). The authors declare that the data supporting the findings of this study are available within the article and supporting materials.

Acknowledgments: MAFF GenBank is acknowledged for the RGMoV isolate. Diffraction data have been collected on BL14.1 beamline at the BESSY II electron storage ring operated by the Helmholtz-Zentrum, Berlin.

Conflicts of Interest: The authors declare that they have no competing interest. The funders of the study had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Toriyama, S.; Mikoshiba, Y.; Doi, Y. Ryegrass mottle virus, a new virus from *Lolium multiflorum* in Japan. *Jpn. J. Phytopathol.* **1983**, *49*, 610–618. [\[CrossRef\]](#)
2. Sarmiento, C.; Sömera, M.; Truve, E. Solemoviruses (Solemoviridae). In *Reference Module in Life Sciences*; Elsevier: Amsterdam, The Netherlands, 2020.
3. Guevara, F.; Flores, F. Near-complete genome sequence of ryegrass mottle virus from irrigation water in Ecuador. *Microbiol. Resour. Announc.* **2021**, *10*, e00037–21. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Balke, I.; Resēviča, G.; Zeltins, A. The ryegrass mottle virus genome codes for a sobemovirus 3C-like serine protease and RNA-dependent RNA polymerase translated via −1 ribosomal frameshifting. *Virus Genes* **2007**, *35*, 395–398. [\[CrossRef\]](#)
5. Ling, R.; Pate, A.E.; Carr, J.P.; Firth, A.E. An essential fifth coding ORF in the sobemoviruses. *Virology* **2013**, *446*, 397–408. [\[CrossRef\]](#)
6. Meier, M.; Truve, E. Sobemoviruses possess a common CfMV-like genomic organization. *Arch. Virol.* **2007**, *152*, 635–640. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Ghosh, A.; Dasgupta, R.; Salerno-Rife, T.; Rutgers, T.; Kaesberg, P. Southern bean mosaic viral RNA has a 5′-linked protein but lacks 3′ terminal poly(A). *Nucleic Acids Res.* **1979**, *7*, 2137–2146. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Olspert, A.; Arike, L.; Peil, L.; Truve, E. Sobemovirus RNA linked to VPg over a threonine residue. *FEBS Lett.* **2011**, *585*, 2979–2985. [\[CrossRef\]](#)
9. Gorbalenya, A.E.; Koonin, E.V.; Blinov, V.M.; Donchenko, A.P. Sobemovirus genome appears to encode a serine protease related to cysteine proteases of picornaviruses. *FEBS Lett.* **1988**, *236*, 287–290. [\[CrossRef\]](#)
10. Nair, S.; Savithri, H.S. Natively unfolded nucleic acid binding P8 domain of SeMV polypeptide 2a affects the novel ATPase activity of the preceding P10 domain. *FEBS Lett.* **2010**, *584*, 571–576. [\[CrossRef\]](#)

11. Sömera, M.; Sarmiento, C.; Truve, E. Overview on sobemoviruses and a proposal for the creation of the family *Sobemoviridae*. *Viruses* **2015**, *7*, 3076–3115. [[CrossRef](#)]
12. Satheshkumar, P.S.; Lokesh, G.L.; Savithri, H.S. Polyprotein processing: *Cis* and *trans* proteolytic activities of Sesbania mosaic virus serine protease. *Virology* **2004**, *318*, 429–438. [[CrossRef](#)] [[PubMed](#)]
13. Olspert, A.; Peil, L.; Hebrard, E.; Fargette, D.; Truve, E. Protein-RNA linkage and post-translational modifications of two sobemovirus VPgs. *J. Gen. Virol.* **2011**, *92*, 445–452. [[CrossRef](#)]
14. Satheshkumar, P.S.; Gayathri, P.; Prasad, K.; Savithri, H.S. “Natively unfolded” VPg is essential for Sesbania mosaic virus serine protease activity. *J. Biol. Chem.* **2005**, *280*, 30291–30300. [[CrossRef](#)] [[PubMed](#)]
15. Nair, S.; Gayathri, P.; Murthy, M.R.; Savithri, H.S. Stacking interactions of W271 and H275 of SeMV serine protease with W43 of natively unfolded VPg confer catalytic activity to protease. *Virology* **2008**, *382*, 83–90. [[CrossRef](#)] [[PubMed](#)]
16. Hebrard, E.; Bessin, Y.; Michon, T.; Longhi, S.; Uversky, V.N.; Delalande, F.; Van Dorsselaer, A.; Romero, P.; Walter, J.; Declerk, N.; et al. Intrinsic disorder in viral proteins genome-linked: Experimental and predictive analyses. *Virol. J.* **2009**, *6*, 23. [[CrossRef](#)] [[PubMed](#)]
17. Balke, I.; Resēviča, G.; Skrastiņa, D.; Zeltiņš, A. Expression and characterisation of the ryegrass mottle virus non-structural proteins. *Latv. Acad. Sciences. Sect. B. Nat. Exact Appl. Sci.* **2010**, *64*, 215–222. [[CrossRef](#)]
18. Grzela, R.; Szolajska, E.; Ebel, C.; Madern, D.; Favier, A.; Wojtal, I.; Zagorski, W.; Chroboczek, J. Virulence factor of potato virus Y, genome-attached terminal protein VPg, is a highly disordered protein. *J. Biol. Chem.* **2008**, *283*, 213–221. [[CrossRef](#)]
19. Pinel-Galzi, A.; Rakotomalala, M.; Sangu, E.; Sorho, F.; Kanyeka, Z.; Traore, O.; Sereme, D.; Poulicard, N.; Rabenantoandro, Y.; Sere, Y.; et al. Theme and variations in the evolutionary pathways to virulence of an RNA plant virus species. *PLoS Pathog.* **2007**, *3*, e180. [[CrossRef](#)]
20. Dixit, K.; Karanth, N.M.; Nair, S.; Kumari, K.; Chakrabarti, K.S.; Savithri, H.S.; Sarma, S.P. Aromatic interactions drive the coupled folding and binding of the intrinsically disordered Sesbania mosaic virus VPg protein. *Biochemistry* **2020**, *59*, 4663–4680. [[CrossRef](#)]
21. Rantalainen, K.I.; Uversky, V.N.; Permi, P.; Kalkkinen, N.; Dunker, A.K.; Makinen, K. Potato virus a genome-linked protein VPg is an intrinsically disordered molten globule-like protein with a hydrophobic core. *Virology* **2008**, *377*, 280–288. [[CrossRef](#)]
22. van de Ven, F.J.; Lycksell, P.O.; van Kammen, A.; Hilbers, C.W. Computer-aided assignment of the 1H-NMR spectrum of the viral-protein-genome-linked polypeptide from cowpea mosaic virus. *Eur. J. Biochem.* **1990**, *190*, 583–591. [[PubMed](#)]
23. Ferrer-Orta, C.; Arias, A.; Agudo, R.; Perez-Luque, R.; Escarmis, C.; Domingo, E.; Verdaguer, N. The structure of a protein primer—polymerase complex in the initiation of genome replication. *EMBO J.* **2006**, *25*, 880–888. [[CrossRef](#)] [[PubMed](#)]
24. Schein, C.H.; Oezguen, N.; Volk, D.E.; Garimella, R.; Paul, A.; Braun, W. NMR structure of the viral peptide linked to the genome (VPg) of poliovirus. *Peptides* **2006**, *27*, 1676–1684. [[CrossRef](#)]
25. Chen, C.; Wang, Y.; Shan, C.; Sun, Y.; Xu, P.; Zhou, H.; Yang, C.; Shi, P.Y.; Rao, Z.; Zhang, B.; et al. Crystal structure of enterovirus 71 RNA-dependent RNA polymerase complexed with its protein primer VPg: Implication for a *trans* mechanism of VPg uridylylation. *J. Virol.* **2013**, *87*, 5755–5768. [[CrossRef](#)] [[PubMed](#)]
26. Gruez, A.; Selisko, B.; Roberts, M.; Bricogne, G.; Bussetta, C.; Jabafi, I.; Coutard, B.; De Palma, A.M.; Neyts, J.; Canard, B. The crystal structure of coxsackievirus B3 RNA-dependent RNA polymerase in complex with its protein primer VPg confirms the existence of a second VPg binding site on *Picornaviridae* polymerases. *J. Virol.* **2008**, *82*, 9577–9590. [[CrossRef](#)]
27. Leen, E.N.; Kwok, K.Y.; Birtley, J.R.; Simpson, P.J.; Subba-Reddy, C.V.; Chaudhry, Y.; Sosnovtsev, S.V.; Green, K.Y.; Prater, S.N.; Tong, M.; et al. Structures of the compact helical core domains of feline calicivirus and murine norovirus VPg proteins. *J. Virol.* **2013**, *87*, 5318–5330. [[CrossRef](#)] [[PubMed](#)]
28. Lee, J.H.; Park, B.S.; Han, K.R.; Biering, S.B.; Kim, S.J.; Choi, J.; Seok, J.H.; Alam, I.; Chung, M.S.; Kim, H.M.; et al. Insight into the interaction between RNA polymerase and VPg for murine norovirus replication. *Front. Microbiol.* **2018**, *9*, 1466. [[CrossRef](#)]
29. Demidiuk, I.V.; Kostrov, S.V. Features of the functional organization of glutamyl endopeptidase. *Mol. Biol.* **1999**, *33*, 100–105.
30. Mil’gotina, E.I.; Voiushina, T.L.; Chestukhina, G.G. Glutamyl endopeptidase. Structure, function, practical use. *Bioorg. Khim.* **2003**, *29*, 563–576.
31. Demidyuk, I.V.; Chukhontseva, K.N.; Kostrov, S.V. Glutamyl endopeptidases: The puzzle of substrate specificity. *Acta Naturae* **2017**, *9*, 17–33. [[CrossRef](#)]
32. Gayathri, P.; Satheshkumar, P.S.; Prasad, K.; Nair, S.; Savithri, H.S.; Murthy, M.R. Crystal structure of the serine protease domain of Sesbania mosaic virus polyprotein and mutational analysis of residues forming the S1-binding pocket. *Virology* **2006**, *346*, 440–451. [[CrossRef](#)] [[PubMed](#)]
33. Koonin, E.V. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* **1991**, *72*, 2197–2206. [[CrossRef](#)]
34. Somera, M.; Truve, E. The genome organization of lucerne transient streak and turnip rosette sobemoviruses revisited. *Arch. Virol.* **2013**, *158*, 673–678. [[CrossRef](#)]
35. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [[CrossRef](#)] [[PubMed](#)]
36. Hirel, P.H.; Schmitter, M.J.; Dessen, P.; Fayat, G.; Blanquet, S. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 8247–8251. [[CrossRef](#)] [[PubMed](#)]

37. Freivalds, J.; Dislers, A.; Ose, V.; Skrastina, D.; Cielens, I.; Pumpens, P.; Sasnauskas, K.; Kazaks, A. Assembly of bacteriophage Qbeta virus-like particles in yeast *Saccharomyces cerevisiae* and *Pichia pastoris*. *J. Biotechnol.* **2006**, *123*, 297–303. [\[CrossRef\]](#)
38. Balke, I.; Resevica, G.; Zeltins, A. Isolation and characterization of two distinct types of unmodified spherical plant sobemovirus-like particles for diagnostic and technical uses. In *Virus-Derived Nanoparticles for Advanced Technologies Methods and Protocols*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 1776, pp. 19–34.
39. Makinen, K.; Makelainen, K.; Arshava, N.; Tamm, T.; Merits, A.; Truve, E.; Zavriev, S.; Saarma, M. Characterization of VPg and the polyprotein processing of cocksfoot mottle virus (genus *Sobemovirus*). *J. Gen. Virol.* **2000**, *81*, 2783–2789. [\[CrossRef\]](#)
40. Receveur-Brechot, V.; Bourhis, J.M.; Uversky, V.N.; Canard, B.; Longhi, S. Assessing protein disorder and induced folding. *Proteins: Struct. Funct. Bioinform.* **2006**, *62*, 24–45. [\[CrossRef\]](#)
41. Nair, S.; Savithri, H.S. Processing of SeMV polyproteins revisited. *Virology* **2010**, *396*, 106–117. [\[CrossRef\]](#)
42. van der Wilk, F.; Verbeek, M.; Dulleman, A.; van den Heuvel, J. The genome-linked protein (VPg) of southern bean mosaic virus is encoded by the ORF2. *Virus Genes* **1998**, *17*, 21–24. [\[CrossRef\]](#)
43. Kader, M.A.; Lindberg, S. Cytosolic calcium and pH signaling in plants under salinity stress. *Plant Signal. Behav.* **2010**, *5*, 233–238. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Plevka, P.; Tars, K.; Zeltins, A.; Balke, I.; Truve, E.; Liljas, L. The three-dimensional structure of ryegrass mottle virus at 2.9 Å resolution. *Virology* **2007**, *369*, 364–374. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Cabot, C.; Martos, S.; Llugany, M.; Gallego, B.; Tolrà, R.; Poschenrieder, C. A role for zinc in plant defense against pathogens and herbivores. *Front. Plant Sci.* **2019**, *10*, 1171. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Gillet, F.X.; Cattoni, D.I.; Petiot-Bécard, S.; Delalande, F.; Poignavent, V.; Brizard, J.P.; Bessin, Y.; Dorsselaer, A.V.; Declerck, N.; Sanglier-Cianférani, S.; et al. The RYMV-encoded viral suppressor of RNA silencing P1 is a zinc-binding protein with redox-dependent flexibility. *J. Mol. Biol.* **2013**, *425*, 2423–2435. [\[CrossRef\]](#)
47. Lacombe, S.; Bangratz, M.; Vignols, F.; Brugidou, C. The rice yellow mottle virus P1 protein exhibits dual functions to suppress and activate gene silencing. *Plant J.* **2010**, *61*, 371–382. [\[CrossRef\]](#)
48. Bonneau, C.; Brugidou, C.; Chen, L.; Beachy, R.N.; Fauquet, C. Expression of the rice yellow mottle virus P1 protein in vitro and in vivo and its involvement in virus spread. *Virology* **1998**, *244*, 79–86. [\[CrossRef\]](#)
49. Poignavent, V.; Hoh, F.; Terral, G.; Yang, Y.; Gillet, F.-X.; Kim, J.-H.; Allemand, F.; Lacombe, E.; Brugidou, C.; Cianferani, S.; et al. A flexible and original architecture of two unrelated zinc fingers underlies the role of the multitask P1 in RYMV spread. *J. Mol. Biol.* **2022**, *434*, 167715. [\[CrossRef\]](#)
50. Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **1999**, *292*, 195–202. [\[CrossRef\]](#)
51. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [\[CrossRef\]](#)
52. Mueller, U.; Förster, R.; Hellmig, M.; Huschmann, F.U.; Kastner, A.; Malecki, P.; Pühringer, S.; Röwer, M.; Sparta, K.; Steffien, M.; et al. The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur. Phys. J. Plus* **2015**, *130*, 141. [\[CrossRef\]](#)
53. Battye, T.G.; Kontogiannis, L.; Johnson, O.; Powell, H.R.; Leslie, A.G. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 271–281. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62*, 72–82. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242. [\[CrossRef\]](#)
56. Vagin, A.; Teplyakov, A. MOLREP: An automated program for molecular replacement. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025. [\[CrossRef\]](#)
57. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. Swiss-model: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Struct. Biol. Crystallogr.* **2004**, *D60*, 2126–2132. [\[CrossRef\]](#)
59. Murshudov, G.N.; Vagin, A.A.; Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53*, 240–255. [\[CrossRef\]](#)
60. Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 352–367. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.