Virology 396 (2010) 106–117

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

# Processing of SeMV polyproteins revisited

Smita Nair, H.S. Savithri *

*Department of Biochemistry, Indian Institute of Science, Bangalore 560012, Karnataka State, India*

ARTICLE INFO

*Article history:*
Received 9 August 2009
Returned to author for revision
1 September 2009
Accepted 19 September 2009
Available online 25 October 2009

*Keywords:*
Sobemoviruses
*Sesbania mosaic virus*
Polyprotein processing
Glutamate specific protease
VPg
P10
P8
*Cis*/*trans* cleavages

ABSTRACT

Processing of *Sesbania mosaic virus* (SeMV) polyprotein 2a and 2ab was reanalyzed in the view of the new genome organization of sobemoviruses. Polyprotein 2a when expressed in *E. coli*, from the new cDNA clone, got cleaved at the earlier identified sites E325-T326, E402-T403 and E498-S499 to release protease, VPg, P10 and P8, respectively. Additionally, a novel cleavage was identified within the protease domain at position E132-S133, which was found to be essential for efficient polyprotein processing. Products, corresponding to cleavages identified in *E. coli*, were also detected in infected Sesbania leaves. Interestingly, though the sites are exactly the same in polyprotein 2ab, it got cleaved between Protease-VPg but not between VPg-RdRp. This indicates to a differential cleavage preference, governed probably by the conformation of 2ab. Also, the studies revealed that, in SeMV, processing is regulated by mode of cleavage and context of the cleavage site.


## Introduction

Sobemoviruses are positive sense ssRNA viruses, named after their type member *Southern bean mosaic virus* and are known to infect both monocots and dicots. These viruses are icosahedral particles of about 30 nm in diameter containing a genomic RNA (gRNA) of approximately 4–4.5 kb in size. The gRNA is covalently attached at its 5′ end to viral protein genome linked (VPg) and lacks a poly A tail at the 3′ end. The gRNA includes four overlapping ORFs. ORFs proximal to the 5′ and 3′ end code for movement protein and coat protein, respectively, and the central two ORFs are translated into polyproteins. Based on their genome organization, sobemoviruses were earlier thought to be either of the Cocksfoot mottle virus (CfMV) or the Southern cowpea mosaic virus (SCPMV) type (Tamm and Truve, 2000). According to this classification, the two types of organization differed only in their central two ORFs. In the CfMV type, ORF2a would code for polyprotein 2a (Protease-VPg) although the domains beyond VPg have not been identified (Fig. 1a). ORF2b would code for RNA dependent RNA polymerase (RdRp) (Fig. 1a). However, in the SCPMV type, the long central ORF2 was proposed to code for polyprotein 2 (Protease-VPg-RdRp) and the small ORF3 (nested within the ORF2) for P3 protein (Fig. 1b). In both types of genome organization, the region immediately upstream of ORF2b/3 had a ribosome frameshift signal, causing the ribosome translating the ORF2a/2 to slip one nucleotide back and translate the ORF2b/3 in −1 frame, resulting either in polyprotein Protease-VPg-RdRP or Protease-VPg-P3 (Lucchesi et al., 2000; Makinen et al., 1995a). *Sesbania mosaic virus* (SeMV) belongs to the Sobemovirus genus and its genome organization was shown to be like that of the SCPMV type (Lokesh et al., 2001).

However, recent sequencing data of 14 strains of RYMV and 4 other sobemoviruses revealed errors in the already reported sequences, showing that all the known sobemoviruses had a common CfMV-like genome organization unlike what was thought earlier (Meier and Truve, 2007). When a new full length cDNA clone of SeMV was sequenced, it revealed that a 'C' was absent from position 2178 (Lokesh et al., 2006) of the earlier reported sequence (Lokesh et al., 2001). In the absence of this 'C', the organization of SeMV gRNA would also be like that of the CfMV type (Fig. 1c). This would imply that in SeMV the central two ORFs will be translated to give two polyproteins, 2a (Protease-VPg-P10-P8) and 2ab (Protease-VPg-RdRp) (Fig. 1c), the C-terminus of 2a and N-terminus of RdRp being different from what was reported previously. Therefore, in the light of the new genome organization for SeMV, the mechanism of processing of polyprotein 2a and 2ab needs to be revisited.

Polyprotein processing is a common strategy employed by many viruses to generate a large number of functional domains from a long protein (Wellink and van Kammen, 1988). Sobemoviral gRNA is known to code for two such polyproteins having a serine protease domain at its N-terminus. To date, the complete genome sequence of

*Abbreviations:* SeMV, *Sesbania mosaic virus*; ssRNA, single-stranded RNA; ORF, open reading frame; VPg, viral protein genome linked; RdRp, RNA dependent RNA polymerase; CfMV, Cocksfoot mottle virus; SCPMV, Southern cowpea mosaic virus; RYMV, Rice yellow mottle virus; P10, protein of size 10 kDa; P8, protein of size 8 kDa.

* Corresponding author. Fax: +91 80 23600814.
*E-mail address:* bchss@biochem.iisc.ernet.in (H.S. Savithri).

0042-6822/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2009.09.025

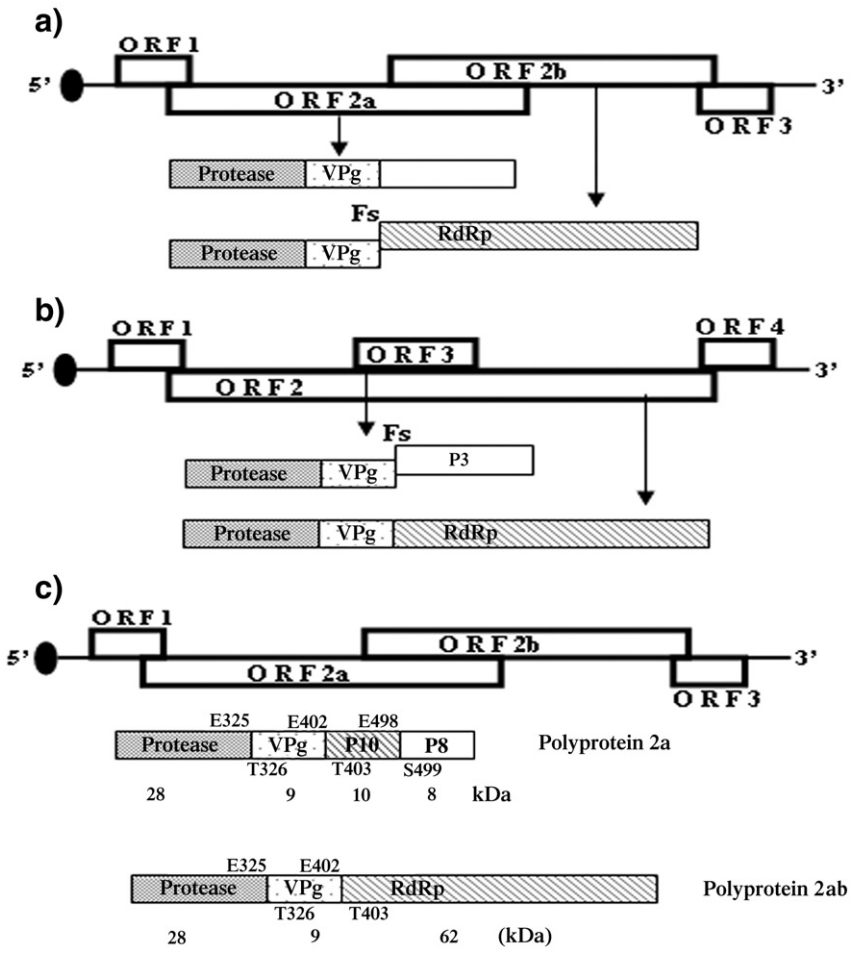


**Fig. 1.** Genome organization of sobemoviruses. Schematic representation of the (a) CfMV type and (b) SCPMV type of genome organization of sobemoviruses. The translation products of the central ORFs are also shown. Fs corresponds to the region where −1 ribosomal frameshift would occur. The domains expected to be translated via −1 ribosomal frameshifting are represented by blocks shifted upwards. (c) Schematic representation of SeMV genome organization and the domain arrangement of Polyprotein 2a and polyprotein 2ab showing the E-T/S cleavage sites and the expected molecular size of each domain.

10 sobemoviruses and 2 unclassified sobemoviruses have been determined (Callaway and Lommel, 2002; Dwyer et al., 2003; Jeffries et al., 1995; Lee and Anderson, 1998; Lokesh et al., 2001; Makinen et al., 1995b; McGavin et al., 2008; Sereme et al., 2008; Wu et al., 1987; Yassi et al., 1994; Zhang et al., 2001; Ziegler et al., 2008). However, the information on the mechanism of polyprotein processing in these viruses, until now, remains sparse. In vitro translation experiments in either rabbit reticulocyte lysate or wheat germ extract have been carried out for six of the sobemoviruses demonstrating the presence of ~100 , ~70 , ~30 (coat protein) and ~18 kDa proteins (ORF1 product) (Tamm and Truve, 2000). The ~100 and ~70 kDa were assigned to the two polyproteins. However, limited polyprotein processing was observed with these in vitro translated products. In CfMV, the cleavage site between protease domain and VPg was identified based on the N-terminal amino acid sequencing of VPg. Several cleavage products were observed with the infected leaf sample (Makinen et al., 2000) and the cleavage sites were only predicted based on the cleavage site sequence at protease-VPg. Only in SeMV, the *E. coli* expressed polyprotein, Protease-VPg-RdRp was shown to undergo processing at E325-T326, E402-T403 and E498-S499 releasing protease VPg, P10 and RdRp domains, respectively (Satheshkumar et al., 2004). But the cleavage at E498-S499 between P10-RdRp was found to be sub optimal and occurred only when the cleavages at other sites were abolished.

In this paper, we demonstrate that SeMV polyprotein 2a expressed from a new cDNA clone (whose sequence was reconfirmed) gets processed at the same E-T/S sites as reported earlier. Furthermore, a novel cleavage site was identified at position E132-S133 that would remove the N-terminal membrane anchoring domain from the protease. Also, both E132-S133 and E498-S499 sites could not be cleaved in *trans* indicating towards their *cis* (intramolecular) recognition. The western analysis of proteins from the SeMV infected leaf samples confirmed that cleavages identified in *E. coli* were also functional *in planta*. Moreover, when polyprotein 2ab (protease-VPg-RdRp) was expressed in *E. coli*, no cleavage was observed to occur between VPg-RdRp suggesting a differential pattern of cleavage for the two polyproteins, 2a and 2ab.

## Results

### *Expression of P8 domain*

The region corresponding to P8 was cloned (Fig. 2a), over-expressed and purified using Ni-NTA affinity chromatography as described in Materials and methods. The purified His-tagged protein moved abnormally as 15 kDa on the SDS-PAGE (Fig. 2b). However, the mass spectrometric analysis of the purified protein showed that it had a molecular mass of 9.766 kDa as is expected for a His-tagged P8 (Fig. 2c). P8 is highly positively charged (theoretical pI-11.75) (Fig. 2d), which could possibly explain its anomalous behavior on the SDS-PAGE. Also, regions corresponding to P10-P8 and VPg-P10-P8 were cloned and expressed. These His-tagged proteins too showed mass abnormality on the SDS-PAGE (data not shown). The purified P8