J. Plant Biochemistry & Biotechnology Vol. 17(2), 201-204, July 2008

Short Communication

Sequence Motifs Comparisons Establish a Functional Portrait of a Multifunctional Protein HC-Pro from Papaya Ringspot Potyvirus

Satendra K Mangrauthia[1], R K Jain[2] and Shelly Praveen[3]*
[1]Division of Biochemistry, [2, 3] Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi 110 012, India

Helper component proteinase (HC-Pro) is a multifunctional protein responsible for multiple molecular events in viral cycle. Here, we demonstrate that functional correlation of sequence motifs of HC-Pro is an important source to predict its role in deubiqutinylation pathway and rescuing viral proteins from degradation. The sequence of papaya ringspot viral HC-Pro was compared with respect to both inter and intra-species across different potyviruses. This study suggested that highly conserved domains involved in post transcriptional gene silencing (PTGS) suppression and proteolytic activity are essential functions in plant-virus cycle. In contrast, mechanisms primed for differentiation such as host specificity and virus replication are less conserved. Also, they contribute substantially to the differences among HC-Pro, derived from different potyviruses. The results obtained from this study provide a framework for new hypothesis and research directions in the area of differential role of potyviral HC-Pro.

Key words: helper component proteinase, PTGS suppressor, potyvirus, proteosome, ubiquitinylation.

Papaya ringspot virus belongs to the family *Potyvirus*, having RNA as genome, that is translated into a polyprotein. This polyprotein is further processed by three virus-encoded proteinases (1), one of these is helper component proteinase (HC-Pro), a multifunctional protein (2). As a strictly *cis-acting* proteinase, it is responsible for its self-cleavage from the polyprotein precursor. It is also involved in a number of infectious processes varying from aphid transmission (3), cell-to-cell long-distance movement (4), genome amplification (5) and suppression of gene silencing mediated host defense (6). HC-Pro also interacts with various host proteins such as calmodulin related protein involved in gene silencing (7). Besides its role in post transcriptional gene silencing (PTGS), its involvement in inhibiting 20S RNase activity of protease complex (proteasome), is reminiscent of its ability in countering the host defense (8). Proteasomes degrade the proteins marked for destruction by attachment of multiple ubiquitin molecules. Its probable role in deubiquitylation activity might be playing a role in counter defense mechanism. Comparative genomics has been a successful tool in identifying functional modules conserved throughout the evolution. The cross species protein domain conservation and variation among HC-Pro domains helps in identification of its structural relatives.

In the present investigations, we have systematically compared and analyzed the sequence domains of HC-Pro involved in different molecular events. The data and analysis resulting from this study provide a framework for new hypothesis and research directions in the area of an interface between viral protein and host machinery.

Maintenance of the Papaya ring spot virus (New Delhi isolate) was done on papaya seedlings through sap inoculation. Total RNA from infected leaves was isolated using RNeasy kit (Qiagen) as per manufacturer's protocol. Total RNA was reverse transcribed using sequence specific reverse primer. Complementary DNA was subjected to PCR amplification using specific primers to amplify the 1.9 Kb region of PRSV genome consisting of 1371 nucleotides representing HC-Pro. Sequence of the forward primer (5′ TGA TGG TAG ATC AAA ACT GGC 3′) was based on the sequence of PRSV genomes available in NCBI database (Acc number: X67673; AY231130; X97251; AY027810; AY162218; AY01072), while the reverse primer was based on the primers used by Charoenslip *et al* (9). The amplicon (~1.9 Kb) comprising 1371 bp of HC-Pro gene along with the flanking regions was subsequently cloned in pGEM-T vector (Promega) and transformed in *Escherichia coli* DH5α. Sequencing was performed at the commercial

*Corresponding author. E-mail : shellypraveen@iari.res.in
1&3 contributed equally.
Abbreviations: PRSV, Papaya ringspot virus; HC-Pro, Helper component proteinase; PTGS, Post transcriptional gene silencing.

Table 1. Source of HC-Pro sequences used in the study from different potyviruses and PRSV isolates

Virus	Acc. Number
Potyviruses	
Lettuce mosaic virus	NP734154
Turnip mosaic virus	NP734214
Plum pox virus	NP734340
Potato virus A	CAA74553
Sweet potato feathery mottle virus	NP734310
Japanese yam mosaic virus	NP734224
Lily mottle virus	NP945137
Scallion mosaic virus	NP734124
Potato virus Y	AAC54827
Tobacco vein mottling virus	NP734329
Peru tomato mosaic virus	NP787939
Konjak mosaic virus	YP529491
Yam mosaic virus	YP022753
PRSV isolates	
New Delhi (from this study)	DQ855428
Brazil	ABD23971
Brazil	ABD23970
Taiwan	NP734234
Taiwan	NP056758
Taiwan	X67673
Thailand-P	AAO16605
Thailand-W	AAG47346

facility using the primers from T7 and SP6 promoter present in the vector.

BioEdit sequence alignment editor version 5.09.04 (10) was used for the analysis of amino acid sequence data. The amino acid sequence of HC-Pro of PRSV (New Delhi isolate, accession number DQ855428) was compared with the corresponding proteins from different potyviruses and with the other PRSV isolates (Table 1) available in NCBI database. Alignment of the HC-Pro proteins was performed by Clustal X version 1.81(11).Gonnet series was followed as protein weight matrix for amino acid alignment. Conserved domain protein architecture of HC-Pro protein was modeled using *All-IN-ONE SEQ-ANALYZER version 1.35* (http://www-personal.umich.edu/~ino/blast.html).

We isolated a gene sequence of 1371 nucleotides (from 1724 to 3094) from Papaya ringspot virus (New Delhi isolate), coding for 457 amino acids having deduced MW of 52 kD and pI of 8.23. A set of sequences from the 5′ terminal of the potyviral genomes taken from NCBI database was compiled and the analysis was restricted to the region from 1724 to 3094 bps relative to the annotated sequences of HC-Pro. Different regions/domains of the HC-Pro were analyzed for sequence-function relationship.

The N-terminal transmission domain was nearly 100 % conserved among all PRSV isolates, whereas only ~18% conserved in comparison with other potyviruses from different hosts. The functional motif KITC [54] is evolutionary conserved in all the potyviruses having binding affinity to the aphid vector stylets. The other conserved motifs like CG[36] and VAAL[41] in all potyviruses may have similar function (Table 2). Beside these, some of the amino acids like H[24], C[26], C[57], F[91], H[93] and L[98] have shown identical positions in all the potyviruses indicating their probable role in metal binding which is supposed to be a key factor in virus transmission (Fig. 1). Cross species protein conservation analysis of the N-terminal region of HC-Pro indicates its close resemblance with the domains possessing affinity for metal binding like the domain of Nif D, a molybdenum-iron protein (Fig. 2).

The central region consists of two RNA binding domains. The first RNA binding domain responsible for genome amplification consists of three conserved motifs among all the potyviruses like FRNK[183], KG[143] and CDNQLD[201]. One unique observation is motif KRT[169] , which is found to be conserved in all the PRSVs, whereas K is replaced by N in all other potyviruses. Conserved domain architectures among different proteins showed its homology with two important protein domains. The RP041 domain having role in the activity of RNA polymerase and Nrap domain is found to be evolutionary conserved from yeast to human, playing crucial role in ribosome biogenesis by interacting with pre rRNA primary transcript (Fig. 2). The second RNA binding domain-having role in PTGS suppression is found to be ~60% conserved. Some of the conserved motifs in this domain are YHAKRFF[219], GY[232], PNG[243] and AIG [250]. This RNA binding domain shares an overlapping functional domain responsible for cell-to-cell movement of the virus. Conserved domain protein architecture reflects close homology of this domain with the domains of membrane binding proteins such as DnaB and Mvi N, suggesting its probable involvement in cell-to-cell movement of virus (Fig. 2).

The proteinase domain of HC-Pro has been mapped at the C-terminal, and 157 amino acids are characterized having cysteine protease like activity. The presence of two conserved amino acid Cys[343] and His[416] at the active site of the protease in all the potyviruses confirmed its probable function uniformly. Beside these two amino acids, other conserved motifs are NIFLAML[352], AELPRILVDH[410], LKANTV[436] and VG[457]. An interesting motif PTK[311] which is found to be evolutionary conserved in all the potyviruses, probably contributes to binding of HC-Pro to the viral coat

Table 2. The sequence motifs identified in helper component proteinase of Papaya ring spot virus (New Delhi isolate)

Motif	Sequence	Location (N-terminal)	Domain	
M1	CG	35-36	Nif D	Transmission
M2	VAAL	38-41	Nif D	
M3	KITC	51-54	Nif D	
M4	LIKG	140-143	RP041 and Nrap	RNA binding
M5	NRT	167-169	RP041 and Nrap	
M5	FRNK	180-183	RP041 and Nrap	
M5	CDNQLD	196-201	—	
M6	NGNF	203-206	—	
M7	WG	208-209	—	
M8	YHAKRFF	213-219	Dna B	
M9	YFE	222-224	Dna B	
M10	GY	231-232	Dna B	
M11	PNG	241-243	Dna B	
M12	AI	248-249	Dna B	
M13	TN	256-257	Dna B	
M14	CCCVT	292-295	—	
M15	PTK	309-311	—	Protease
M16	GN	317-318	—	
M17	GD	320-321	—	
M18	GYCY	341-344	Peptidase C19	
M19	NIFLAML	346—352	Peptidase C19	
M20	AK	360-361	Peptidase C19	
M21	FTK	363-365	Peptidase C19	
M21	VRD	367-369	Peptidase C19	
M22	LG	375-376	Peptidase C19	
M23	WP	378-379	Peptidase C19	
M24	AT	385-387	Peptidase C19	
M25	AELPRILVDH	401-410	Peptidase C19	
M26	HV	416-417	Peptidase C19	
M27	DS	419-420	Peptidase C19	
M28	GS	422-423	Peptidase C19	
M29	TGYH	426-429	Peptidase C19	
M30	LKANTV	431-436	Peptidase C19	
M31	QL	438-439	—	
M32	VG	456-457	—	

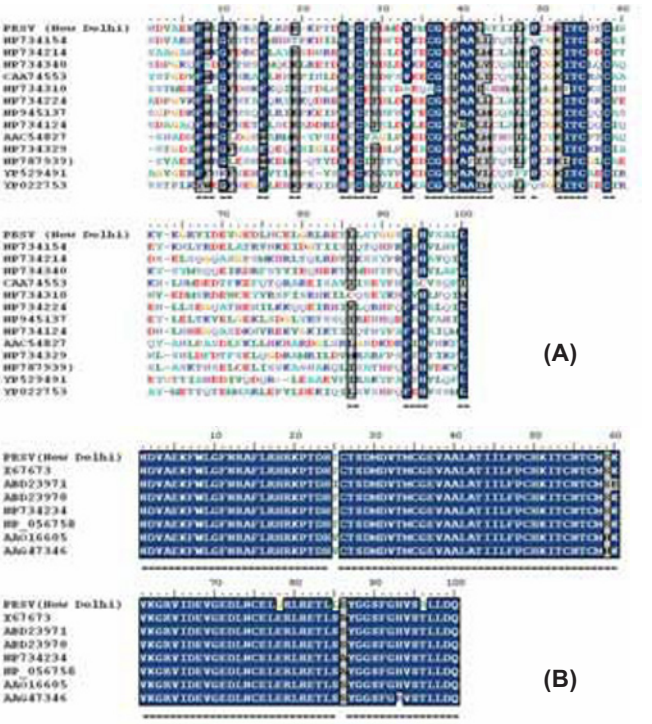

Fig. 1. Multiple sequence alignment of N-terminal region (1-100 amino acids) of HC-Pro. **(A)** Sequence of HC-Pro from different potyviruses, and **(B)** Sequence analysis of HC-Pro from different PRSV isolates. Conserved domains/motifs are shown by dotted lines.

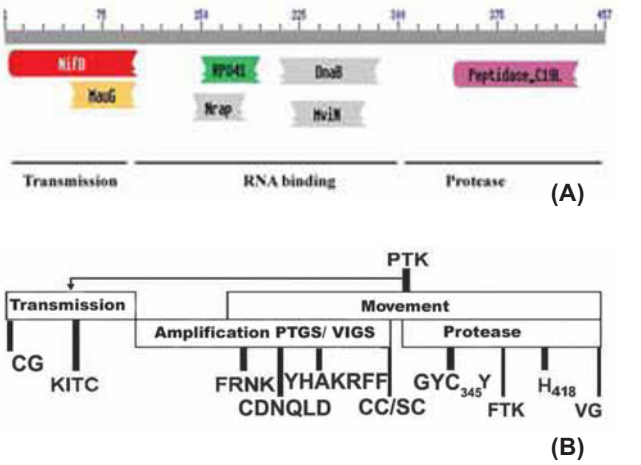

Fig. 2. Cross species conservation at the three functional regions of HC-Pro. **(A)** The image generated by *All-IN-ONE SEQ-ANALYZER* version 1.35, depicting different functional domains of HC-Pro, based on homology in sequence and their functional attributes, and **(B)** Specific conserved sequence motifs in three functional regions of HC-Pro.

protein. The presence of many conserved motifs in this region confirms its fundamental role as proteolytic enzyme in all the potyviruses irrespective of their host. This region shows strong homology with the other peptidases when compared with cross protein conserved domain architecture. Its close homology with the peptidase C19 L, a subfamily of peptidase C19, reflects an additional role of this protease beside autocleavage (Fig. 2). Proteases of this family are involved in intracellular proteolytic activity that removes ubiquitin molecule from polyubiquinated peptides, hence affecting the protein turnover through the proteosome system (12).

Despite the fact that only PRSV-HC-Pro gene was used in this study, a number of functional modules were identified that were conserved and thus predicted to be essential for performing multiple functions. This study generated a conserved domain protein architecture and comprehensive functional portrait of HC-Pro, featured by conserved and divergent landscapes emphasizing

fundamental and species-specific mechanisms. The data and analysis resulting from this study provide a framework for new hypotheseis and research directions for functional genomics of HC-Pro in relation to different hosts.

We demonstrate that sequence motifs comparison is a powerful tool to predict functional mechanisms. Much of our current knowledge of functional domains of HC-Pro is supported or reaffirmed by this correlation analysis. The most interesting N-terminal domain is the first 100 amino acids having a putative Zn finger motif involved in virus transmission (13). Sequence analysis of this region strongly suggests that variability at the N terminal is due to host-virus interaction in different potyviruses. The N-terminal region of the potyviruses (PRSV) from the same host papaya is found to be universally conserved reflecting its direct relationship with the host. Although the presence of conserved motif KITC having interaction with aphid stylets is found universally conserved in all potyviruses transmitted through aphids. Central region of HC-Pro from 101-300 amino acids is assumed to be important in genome amplification as well as PTGS suppression (2, 6). It has two RNA binding domains, which is evident from high lysine, arginine and asparagine content in this region. Probably, one is playing a role in viral RNA binding for genome amplification, which is found to be variable, as evident from its role in binding with diverse viral genomes. While other may be involved in binding with small RNAs to inhibit intermediate step of PTGS, which is size specific rather than sequence specific, hence more conserved (14). The annotation of C-terminal domain with peptidase C-19L having unique property of deubiquitylation suggests its role in rescuing viral proteins from proteolytic cleavage with host proteosome (8, 15). Our analysis proposes a model for the probable function of this protease in deubiquitinylation of viral proteins which is in close agreement of SARS coronaviral PLpro and rescuing them from the degradation in the host proteosome. This suggests one more level of virus counter defense at the protein level (16).

In conclusion, this study provides information on the various sequence motifs of potyviral multifunctional protein HC-Pro to relate its biological functions. We predicted here the basis of host specificity through transmission with the metal binding domains. It is interesting to observe two RNA binding domains being involved in PTGS suppression and genome amplification. The most important finding is to presenting the form of hypothesis that suggests the secondary role of proteinase in rescuing the viral proteins from degradation. The functional portrait of HC-Pro profiled by this study provides a basis for defining the "counter defense strategy" in terms of regulating host proteosomal activities. The findings are also important in advancing our understanding of the role of HC-Pro in plant virus interaction.

Acknowledgements

SKM was supported by a fellowship from Indian Agricultural Research Institute, New Delhi. Laboratory support by Mr CM Kushwaha and financial support by Department of Biotechnology, Govt. of India are acknowledged.

Received 7 December, 2007; accepted 7 July, 2008.

Online published 14 July, 2008.

References

1 **Carrington JC, Freed DD & Oh CS**, *EMBO J*, **9** (1990) 1347.

2 **Varrelmann M, Maiss E, Pilot R & Palkovics L**, *J Gen Virol,***88** (2007) 1005.

3 **Dombrovsky A, Gollop N, Chen S, Chejanovsky A & Raccah B**, *J Gen Virol,* **88** (2007) 1602.

4 **Saenz P, Salvado BR, Simon-Mateo C, Kasschau KD, Carrington JC & Garcia JA**, *J Virol,* **76** (2002) 1922.

5 **Cronin S, Verchot J, Haldeman-Cahill R, Schaad MC & Carrington JC**, *Plant Cell*, **7** (1995) 549.

6 **Llave C, Kasschau KD & Carrington JC**, *Proc Natl Acad Sci,* USA, **97** (2000) 13401.

7 **Anandlakshmi R, Marathe R, Ge X, Herr Jr JM, Mau C, Mallory AC, Pruss GJ, Bowman L & Vance VB**, *Science,* **290 (**2000) 142.

8 **Ballut L, Drucker M, Pugniere M, Cambon F, Blanc S, Roquet F, Candresse T, Schmid HP, Niclas P, Gall OL, & Badaoui S**, *J Gen Virol*,**86** (2005) 2595.

9 **Charoenslip G, Attasart P, Juricek M, Panyim S & Kertbundit S**, *Science Asia,* **29** (2003) 89.

10 **Hall TA**, *Nucl Acids Symp Ser*, **41** (1999) 95.

11 **Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DJ**, *Nucl Acids Res,* **24** (1997) 4876.

12 **Ratia K, Saikatendu S K, Bernard D, Barretto S N, Baker S C, Stevens R C & Mesecar A D,** *Proc Natl Acad Sci,* USA, **103** (2006) 5717**.**

13 **Robaglia C, Durand-Tardif M, Tronchet M, Boudazin G, Astier- Manifacier S & Casse-Delbart F**, *J Gen Virol*, **70** (1989) 935.

14 **Merai Z, Kerenyi Z, Kertesz S, Magna M, Lakatos L & Silhavy D**, *J Virol,* **80** (2006) 5747.

15 **Jin Y, Ma D, Dong J, Jin J, Li D, Deng C & Wang T**, *J Virol,* **81** (2007) 12881.

16 **Sulea T, Lindner GA, Purisima EO & Menard R**, *J Virol,* **79** (2005) 4550.