



# Plant Pathology Journal

ISSN 1812-5387

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

RESEARCH ARTICLE

OPEN ACCESS

DOI: 10.3923/ppj.2015.123.129

## Sequence Motif Comparison and Homology Modeling of Helper Component Proteinase (HC-Pro) of *Banana bract mosaic virus*

Anuradha Chelliah, Balasubramanian Velusamy and Selvarajan Ramasamy  
ICAR-National Research Centre for Banana, Thogamalai Road Thayanur Post Tiruchirapalli, Tamil Nadu, India

### ARTICLE INFO

#### Article History:

Received: April 24, 2015

Accepted: June 10, 2015

#### Corresponding Author:

Selvarajan Ramasamy,  
ICAR-National Research Centre for  
Banana, Thogamalai Road Thayanur  
Post Tiruchirapalli, Tamil Nadu, India  
Tel: 91431-2618106/2618104  
Fax: 91431-2618115/+862885415389

### ABSTRACT

The Helper Component Proteinase (HC-Pro) is a key protein encoded by the plant viruses of the genus potyvirus. The HC-Pro is involved in different steps of the viral cycle, aphid transmission, replication virus cell-to-cell and systemic movement and is a suppressor of post-transcriptional gene silencing. The amino acid sequence of *Banana bract mosaic virus* (BBrMV) HC-Pro was compared both inter and intra species across different potyviruses to demonstrate the functional correlation of sequence motifs of HC-Pro to predict its role in virus cycles. Knowledge on the structure of HC-Pro will greatly contribute to understand its multifaceted functions and how structural domains are organized to fulfill them in concert or independently hence, a stable model of BBrMV HC-Pro has been proposed based on homology modeling. The results obtained from this study will provide a framework for new hypothesis and research directions in the area of differential role of HC-Pro.

**Key words:** *Potyvirus*, helper component proteinase, *Banana bract mosaic virus*, motifs, homology modelling

### INTRODUCTION

*Banana bract mosaic virus* (BBrMV), a member of the genus potyvirus, family Potyviridae, has a single stranded positive sense RNA of about 9.7 kb in length (Balasubramanian and Selvarajan, 2012). *Banana bract mosaic virus* (BBrMV) encodes a single polypeptide that is processed by three viral proteinases to release all viral proteins needed for the infection. One of these proteins is Helper Component Proteinase (HC-Pro), which is involved in many functions during the virus life cycle (Varrelmann *et al.*, 2007). Sequence alignments and motif identification has suggested that HC-Pro can be schematically divided into three regions: An N-terminal region containing cysteine rich region with a zinc finger like metal binding motif responsible for symptom severity, systemic infection and viral transmission by aphids (Govier and Kassanis, 1974). C-terminal region harboring the proteinase activity, catalyses autoproteolytic cleavage from the polyprotein and plays an important role in virus cell to cell movement and a central region is important for genome amplification, synergism with other viruses and systemic movement within the host plant (Kasschau *et al.*, 1997; Urcuqui-Inchima *et al.*, 2000). Structural knowledge of HC-Pro is necessary to understand its multiple functions. In

the present study, *in silico* analysis has been carried out to compare and analyze the sequence domains and motifs of BBrMV HC-Pro involved in different molecular events and a computational approach was used to predict and validate the three dimensional structure of BBrMV HC-Pro. The data and the analysis resulting from this study will help in focusing in the area of interface between viral and host machinery.

### MATERIALS AND METHODS

*Banana bract mosaic virus* (BBrMV-Try) infected leaves were collected from Nendran cv. at National Research Center for Banana farm, Trichy. HC-Pro primers were commercially synthesized: Sequences of the forward (5'-AGCGAGCCAGCAGCAGGCCCA TT CTG-3') and reverse (5'-GCCAACCAATAATCTTTCAT3') primers were based on the sequence of BBrMV-Try available in NCBI database. Total RNA was isolated from infected banana leaves using RNeasy plant minikit (Qiagen, USA) and viral RNA was reverse transcribed using Revert Aid™ H Minus first strand cDNA synthesis kit (MBI Fermentas, USA) using oligo (dT) as primer by following the manufacturer's protocol. Complementary DNA was subjected to PCR amplification using virus specific primers to amplify ~ 1500 bp of

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

Viruses	Size (bp)	GenBank accession No.	References
<b>Potyviruses</b>			
<i>Lettuce mosaic virus</i> N	1374	NP734154	Revers <i>et al.</i> (1997)
<i>Turnip mosaic virus</i>	1374	NP734214	Jenner <i>et al.</i> (2000)
<i>Plum pox virus</i>	1374	NP734340	Maiss <i>et al.</i> (1989)
<i>Sweet potato feathery mottle virus</i>	1374	NP734310	Sakai <i>et al.</i> (1997)
<i>Japanese yam mosaic virus</i>	1374	NP734224	Fuji and Nakamae (2000)
<i>Lily mottle virus</i>	1374	NP945137	Zheng <i>et al.</i> (2003)
<i>Scallion mosaic virus</i>	1371	NP734124	Che <i>et al.</i> (2002)
<i>Potato virus</i>	1368	YAAC54827	Maia and Bernardi (1996)
<i>Peru tomato mosaic virus</i>	1368	NP787939	Spetz <i>et al.</i> (2003)
<i>Konjak mosaic virus</i>	1368	YP529491	Nishiguchi <i>et al.</i> (2006)
<i>Yam mosaic virus</i>	1374	YP022753	Aleman <i>et al.</i> (1996)
<b>BBrMV isolates</b>			
Trichy	1370	HM131454	This study
Philippines	1370	DQ851496	Ha <i>et al.</i> (2008)

HC-Pro gene along with the flanking regions. It was subsequently cloned in pGEM-T vector (Promega) and sequenced. Bioedit sequence alignment editor version 5.09.04 (Hall, 1999) was used for the analysis of amino acid sequence data. The amino acid sequence of BBrMV-Try HC-Pro (Accession number-HM131454) was compared with the corresponding proteins from different potyviruses including a BBrMV-Phillipines isolate available at NCBI database (Table 1). The conserved domain in the predicted HC-Pro protein was explored using NCBI's Conserved Domain Database (CDD) (<http://www.ncbi.nlm.nih.gov/cdd>). The PROSITE was used to scan the protein sequences for identification of motifs (<http://prosite.expasy.org/>). Multiple sequence alignment was done using CLUSTAL W program. The secondary structure of the BBrMV-Try HC-Pro was constructed by homology comparative modeling (<http://swissmodel.expasy.org/>). Candidate structures for homology modeling were selected according to pairwise alignment scores and their similarities in cysteine distribution patterns relative to the target sequences. In this study, *Turnip mosaic virus* HC-Pro (PDB code: 3rNVA) was used as template structure for homology modeling of BBrMV-Try HC-Pro. ESyPred3D web server 1.0 was used for 3D structure generation based on the information obtained from sequence alignment (<http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esyPred/>). ESyPred3D generated structure was further verified by PROCHECK (<http://www.biochem.ulg.ac.uk/~roman/procheck/procheck.html/>), WHATIF (<http://swift.cmbi.ru.nl/whatif/>) and Swiss Model.

## RESULTS AND DISCUSSION

The HC-Pro sequence of BBrMV-Try isolate originating from India was determined, analyzed and compared with a BBrMV isolate originated from Phillipines (Fig. 1a). We found that the sequence length of the HC-Pro gene of BBrMV-Try was 1368 nt and encoding a protein of 456 AA having a deduced molecular weight of 51.88 kDa and pI of 8.39. The size of HC-Pro was the same as that of Phillipines isolate of BBrMV. The nucleotide sequence of BBrMV-Try showed 95.1% identity with BBrMV-Phi isolate compared in this study (data not shown). In the deduced amino acid sequences, BBrMV-Try isolate shared 98% identity with BBrMV-Phi isolate. Different regions and domains of HC-Pro were analysed for sequence and function relationship.

Table 2: Sequence motifs identified in helper component proteinase of *Banana bract mosaic virus* (Trichy isolate)

Motifs	Sequence	Location	Domain
1	CG	35-36	N-Terminal
2	MAQL	39-43	N-Terminal
3	RISC	51-54	N-Terminal
4	FRNK	180-183	Central region
5	CDNQLD	196-201	Central region
6	NGNF	203-206	Central region
7	WN	208-209	Central region
8	HAKRFF	214-218	Central region
9	YFD	222-224	Central region
10	GY	231-232	Central region
11	PNC	241-243	Central region
12	AI	248-249	Central region
13	TN	256-257	Central region
14	CCCTV	291-295	Central region
15	PIY	301-304	C-Terminal
16	GD	320-321	C-Terminal
17	GYCY	341-344	C-Terminal
18	NIFLAML	346-352	C-Terminal
19	AK	360-361	C-Terminal
20	FTV	363-365	C-Terminal
21	LG	375-376	C-Terminal
22	WP	378-379	C-Terminal
23	AELPRILVDH	401-410	C-Terminal
24	HV	416-417	C-Terminal
25	DS	419-420	C-Terminal
26	GS	422-423	C-Terminal
27	TGYH	426-429	C-Terminal
28	LKANTV	431-436	C-Terminal
29	QL	438-439	C-Terminal
30	VG	456-457	C-Terminal

The N-terminal transmission domain of Try isolate was nearly 96% conserved with the Phi isolate. Only ~7-27% was conserved in comparison with other potyviruses from different hosts. Cross species protein conservation analysis of the N-terminal region of BBrMV-Try HC-Pro indicates its close resemblance with the domains possessing affinity for metal binding like the domain of Nif D, a molybdenum iron protein. The functional motif R(K)IS(T)C<sup>55</sup> is evolutionarily conserved in all the potyviruses is involved in binding to the aphid vector's stylets. The other conserved motifs like CG<sup>39</sup> and L(V)M(A)AQ(L)<sup>39</sup> may have similar function (Table 2). Beside these, some of the amino acids like H<sup>27</sup>, C<sup>29</sup>, C<sup>58</sup> and F<sup>96</sup> 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. 1b).

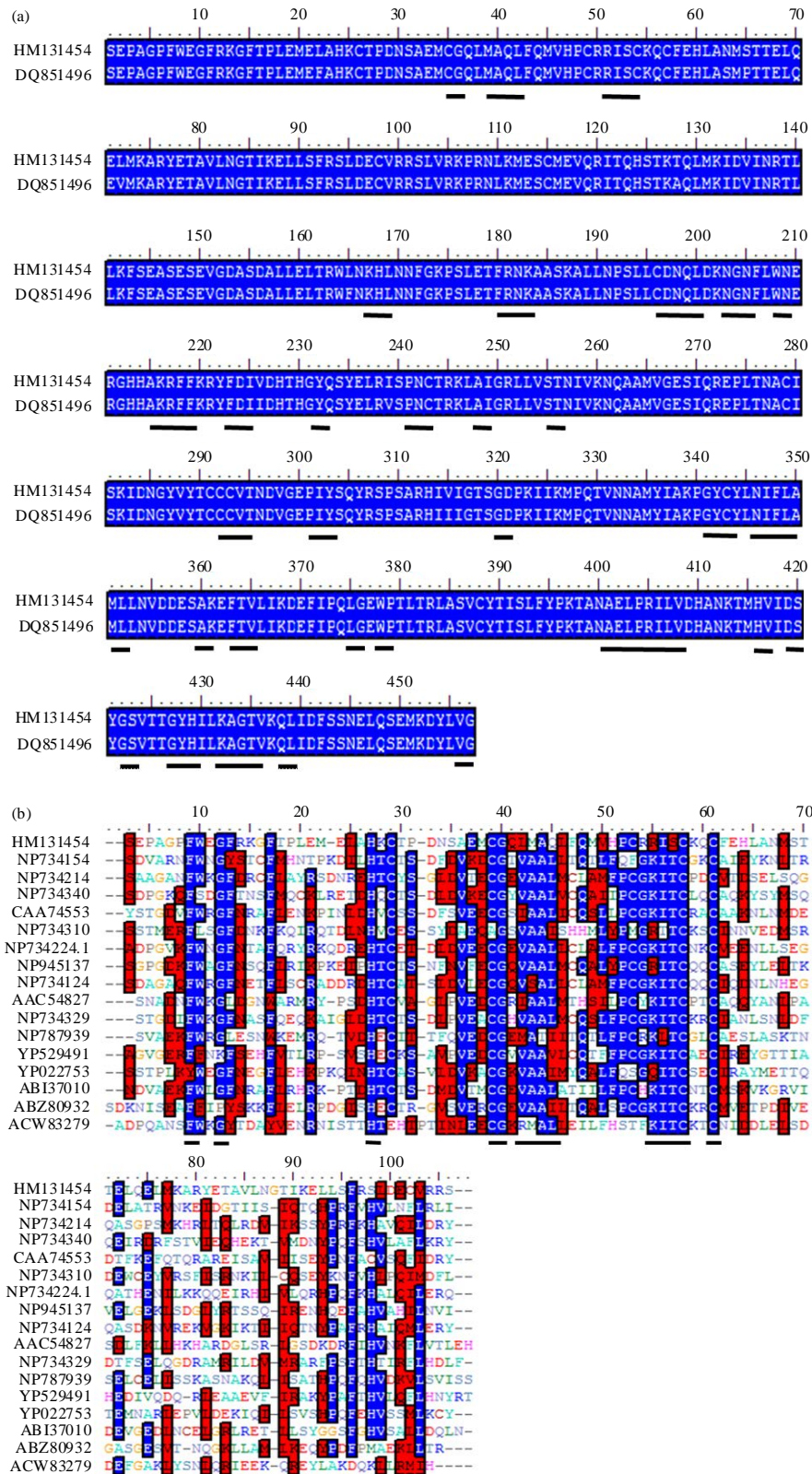


Fig. 1(a-b): Multiple sequence alignment of HC-Pro, (a) Sequence analysis of HC-Pro from different BBrMV isolates and (b) N-terminal region (1-100 amino acids) of HC-Pro from different potyviruses. Conserved domains/motifs are shown by solid lines

The central region of HC-Pro (100-300 AA) 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<sup>197</sup>, KF(G)<sup>153</sup> CDNQLD<sup>214</sup> and IGN(R)<sup>275-277</sup>. The second RNA binding domain, having role in PTGS suppression is also found to be conserved. Some of the conserved motifs in this domain are H(Y) HAKRFF<sup>239</sup>, GY<sup>257</sup>, PNC(G)<sup>267</sup> and AIG<sup>274</sup>. The motif CC<sup>317-318</sup> is involved synergism with other viruses and systemic movement within the host plant (Kasschau *et al.*, 1997; Cronin *et al.*, 1995).

The proteinase domain of HC-Pro has been mapped to the C-terminal 156 amino acids and it shows the presence of two conserved amino acid Cys<sup>369</sup> and His<sup>442</sup> at the active site of the protease in all the potyviruses (Oh *et al.*, 1989) and other conserved motifs are NIFLAML<sup>372</sup>, AELPRILVDH<sup>427</sup>, LKAG(N)TV<sup>457</sup> and VG<sup>482</sup>. The other is the C-terminal PTK motif (310-312 AA) that probably contributes to binding of HC-Pro to the viral capsid protein's N-terminal DAG motif (Blanc *et al.*, 1997; Peng *et al.*, 1998). One unique observation is motif PSA<sup>335</sup>, which is found evolutionary conserved in both the BBrMV isolates, whereas SA is replaced by TK in all other potyviruses. 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. Proteases of this family are involved in intracellular proteolytic activity that removes ubiquitin molecule from polyubiquitinated peptides, hence affecting the protein turnover through the proteasome system (Ratia *et al.*, 2006).

A computational approach following homology modeling protocol has been used to predict the 3D structure of BBrMV-TRY HC-Pro protein as only a low resolution structure for HC-Pro has been obtained by means of 2D crystallography and electron microscopy and it has not yet been determined experimentally (Guo *et al.*, 2011) and this will also help in relating the structure with biological functions. Template identification was carried out by BLASTp search against PDB database in Swiss Work Model/Expasy. Based on the maximum identity with high score (169) and lower e-value ( $3e^{-42}$ ) 3RNV (Chain A) was used as the template for homology modeling and sequence alignment between HC-Pro protein and 3RNV revealed that the first 337 residues were deleted because the corresponding residues were not present in the template and homology modeling was carried out from 335-447 residues of the C-terminal region followed by energy minimization by SwissPdb Viewer. The final stable structure of BBrMV-TRY HC-Pro is shown in Fig. 2a-b. The boundary of the proteinase domain (336-456) defined by the structure is

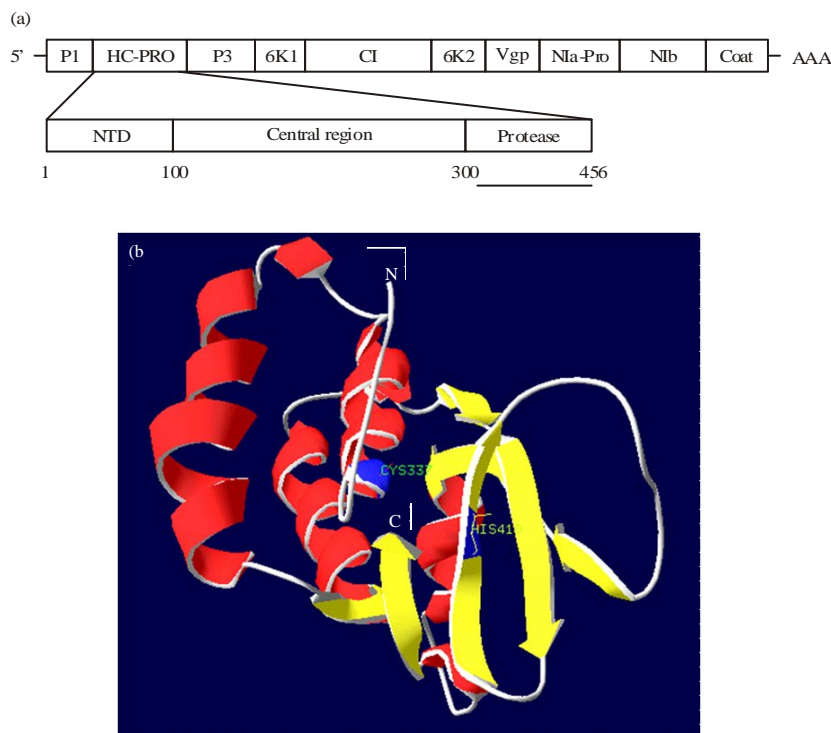


Fig. 2(a-b): Three dimensional structure of BBrMV HC-Pro, (a) Diagram of the BBrMV RNA genome, the domain organization of HC-Pro and the region used for modelling and (b) Ribbon representation of the CPD structure, the N and C terminus are indicated. The catalytic dyad residues Cys337 and His 410 (green) are shown as sticks

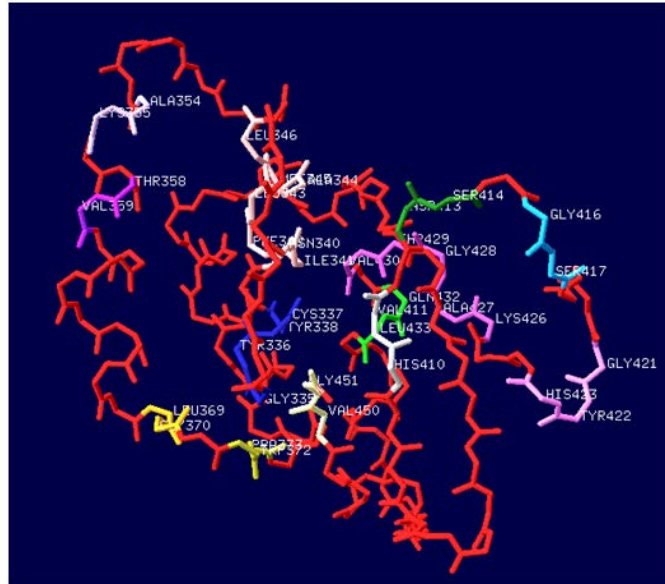


Fig. 3: Sequence motifs in the BBrMV HC-Pro CPD. Specific conserved motifs in the C-terminal domain indicated in different colors

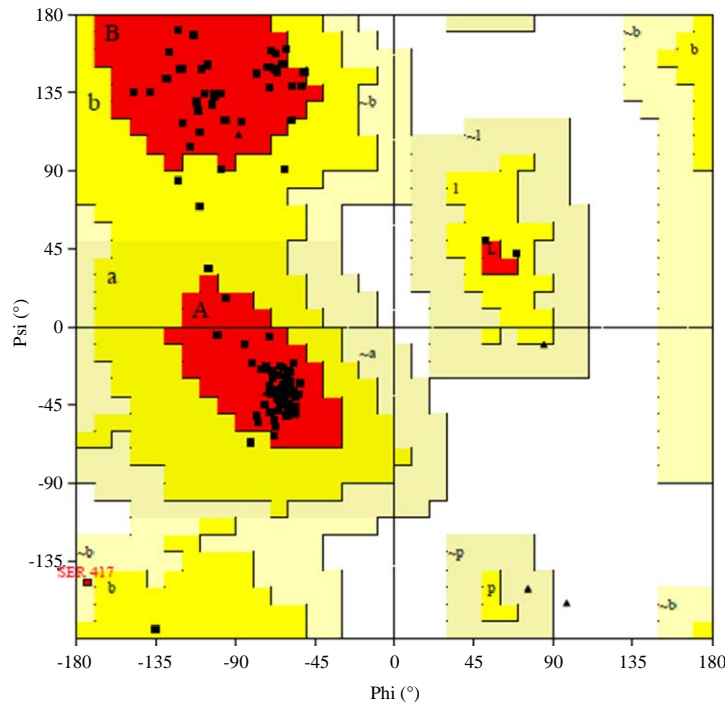


Fig. 4: Ramachandran plot for the BBrMV HC-Pro protein

slightly shorter at its N terminus than previously defined and excludes the PT(S)K(A) motif that is involved in aphid mediated transmission. Figure 3 shows all the conserved motifs present in the C-terminal region and also the Cysteine and Histidine residues that are essential for autoproteolysis.

To verify further the predicted structure, a Ramachandran plot was computed with PROCHECK program

([www.ebi.ac.uk/thornton-srv/software/PROCHECK/](http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/)). It showed that 92.8% residues are in the most favored regions, 6.3% in additional allowed regions, 0.9% in generously allowed regions and 0.0% in the disallowed conformations (Fig. 4). The overall PROCHECK G-factor for the homology modeled structure was 0.12. This score indicates that the modeled structure is acceptable. The

structural superimposition of C $\alpha$  trace of the model over template structure 3RNV was resulting in a Root Mean Square Deviation (RMSD) of 0.17Å (<http://wishart.biology.ualberta.ca/SuperPose>) and the QMEAN6 score for the whole model was 0.655 and the QMEAN Z-Score was -0.922, which reflects the whole model reliability (<http://swissmodel.expasy.org>). The overall quality factor also called ERRAT score was obtained as 98.2, supporting its good quality (<http://nihserver.mbi.ucla.edu/ERRATv2/>). Energetic architecture as revealed by PROSA (data not shown) score was negative (-5.7) for the modeled protein, value quite similar to that was obtained for template (-6.2), which indicates its correctness (<https://prosa.services.came.sbg.ac.at/prosa.php/>). Results of WHATIF also indicate about the fitness of the modeled structure. Based on these results, it was ascertained that obtained structure is of reasonably good quality.

### CONCLUSION

In conclusion, this study provides information on various sequence motifs, domains of BBrMV HC-Pro to relate its biological functions in understanding its role in plant virus interaction. In addition, the homology model of BBrMV HC-Pro provides important insight into its unique folds, catalysis mechanism, substrate specificity and cis acting mode.

### REFERENCES

- Aleman, M.E., J.F. Marcos, C. Brugidou, R.N. Beachy and C. Fauquet, 1996. The complete nucleotide sequence of yam mosaic virus (Ivory Coast isolate) genomic RNA. *Arch. Virol.*, 141: 1259-1278.
- Balasubramanian, V. and R. Selvarajan, 2012. Complete genome sequence of a banana bract mosaic virus isolate infecting the French plantain cv. Nendran in India. *Arch. Virol.*, 157: 397-400.
- Blanc, S., J.J. Lopez-Moya, R.Y. Wang, S. Garcia-Lampasona, D.W. Thornbury and T.P. Pirone, 1997. A specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus. *Virology*, 231: 141-147.
- Che, J., H.Y. Zheng, J.P. Chen and M.J. Adams, 2002. Characterisation of a potyvirus and a potexvirus from Chinese scallion. *Arch. Virol.*, 147: 683-693.
- Cronin, S., J. Verchot, R. Haldeman-Cahill, M.C. Schaad and J.C. Carrington, 1995. Long-distance movement factor: A transport function of the potyvirus helper component proteinase. *Plant Cell*, 7: 549-559.
- Fuji, S. and H. Nakamae, 2000. Complete nucleotide sequence of the genomic RNA of a mild strain of Japanese yam mosaic potyvirus. *Arch. Virol.*, 145: 635-640.
- Govier, D.A. and B. Kassanis, 1974. A virus-induced component of plant sap needed when aphids acquire potato virus Y from purified preparations. *Virology*, 61: 420-426.
- Guo, B., J. Lin and K. Ye, 2011. Structure of the autocatalytic cysteine protease domain of potyvirus helper-component proteinase. *J. Biol. Chem.*, 286: 21937-21943.
- Ha, C., S. Coombs, P.A. Revill, R.M. Harding, M. Vu and J.L. Dale, 2008. Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. *Arch. Virol.*, 153: 25-36.
- Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acid Symp. Ser.*, 41: 95-98.
- Jenner, C.E., F. Sanchez, S.B. Nettleship, G.D. Foster, F. Ponz and J.A. Walsh, 2000. The cylindrical inclusion gene of Turnip mosaic virus encodes a pathogenic determinant to the Brassica resistance gene TuRB01. *Mol. Plant Microbe Interact.*, 13: 1102-1108.
- Kasschau, K.D., S. Cronin and J.C. Carrington, 1997. Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology*, 228: 251-262.
- Maia, I.G. and F. Bernardi, 1996. Nucleic acid-binding properties of a bacterially expressed potato virus Y helper component-proteinase. *J. Gen. Virol.*, 77: 869-877.
- Maiss, E., U. Timpe, A. Briske, W. Jelkmann and R. Casper *et al.*, 1989. The complete nucleotide sequence of plum pox virus RNA. *J. Gen. Virol.*, 70: 513-524.
- Nishiguchi, M., S. Yamasaki, X.Z. Lu, A. Shimoyama and K. Hanada *et al.*, 2006. Konjak mosaic virus: The complete nucleotide sequence of the genomic RNA and its comparison with other potyviruses. *Arch. Virol.*, 151: 1643-1650.
- Oh, C.S. and J.C. Carrington, 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. *Virology*, 173: 692-699.
- Peng, Y.H., D. Kadoury, A. Gal-On, H. Huet, Y. Wang and B. Raccach, 1998. Mutations in the HC-Pro gene of zucchini yellow mosaic potyvirus: Effects on aphid transmission and binding to purified virions. *J. Gen. Virol.*, 79: 897-904.
- Ratia, K., K.S. Saikatendu, B.D. Santarsiero, N. Barretto, S.C. Baker, R.C. Stevens and A.D. Mesecar, 2006. Severe acute respiratory syndrome coronavirus papain-like protease: Structure of a viral deubiquitinating enzyme. *Proc. Natl. Acad. Sci. USA.*, 103: 5717-5722.
- Revers, F., S.J. Yang, J. Walter, S. Souche and H. Lot *et al.*, 1997. Comparison of the complete nucleotide sequences of two isolates of lettuce mosaic virus differing in their biological properties. *Virus Res.*, 47: 167-177.
- Sakai, J., M. Mori, T. Morishita, M. Tanaka, K. Hanada, T. Usugi and M. Nishiguchi, 1997. Complete nucleotide sequence and genome organization of sweet potato feathery mottle virus (S strain) genomic RNA: The large coding region of the P1 gene. *Arch. Virol.*, 142: 1553-1562.
- Spetz, C., A.M. Taboada, S. Darwich, J. Ramsell, L.F. Salazar and J.P. Valkonen, 2003. Molecular resolution of a complex of potyviruses infecting solanaceous crops at the centre of origin in Peru. *J. Gen. Virol.*, 84: 2565-2578.

- Urcuqui-Inchima, S., I.G. Maia, P. Arruda, A.L. Haenni and F. Bernardi, 2000. Deletion mapping of the potyviral helper component-proteinase reveals two regions involved in RNA binding. *Virology*, 268: 104-111.
- Varrelmann, M., E. Maiss, R. Pilot and L. Palkovics, 2007. Use of pentapeptide-insertion scanning mutagenesis for functional mapping of the plum pox virus helper component proteinase suppressor of gene silencing. *J. Gen. Virol.*, 88: 1005-1015.
- Zheng, H.Y., J. Chen, M.F. Zhao, L. Lin, J.P. Chen, J.F. Antoniw and M.J. Adams, 2003. Occurrence and sequences of *Lily mottle virus* and *Lily symptomless virus* in plants grown from imported bulbs in Zhejiang province, China. *Arch. Virol.*, 148: 2419-2428.