Mysore J. Agric. Sci., 58 (1) : 222-231 (2024)

Analysis of Chilli Chloroplastic Photosystem II Protein D1 Interaction with the Cucumber Mosaic Virus (CMV) Coat Protein Using Computational Tools

P. S. Pooja[1], Kopparthi Amruthavalli Sindhura[2], Shadab M. Khatib[3], M. Bharath[4], S. S. Kavya[5], Mantesh Muttappagol[6], Shridhar S. Hiremath[7], K. P. Lavanya[8], T. M. Ningaraju[9] and C. N. Lakshminarayana Reddy[10]

[1, 3, 4, 5, 6, 7, 8 & 10]Department of Plant Pathology, [2]Department of Agricultural Entomology, [9]Department of Plant Biotechnology, College of Agriculture, UAS, GKVK, Bengaluru - 560 065

e-Mail : cnlreddy@gmail.com

Authors Contribution

P. S. Pooja : Conceptualisation & visualization; Kopparthi Amruthavalli Sindhura, Shadab M. Khatib, M. Bharath, S.S. Kavya & K.P. Lavanya : Conceptualisation and methodology; Mantesh Muttappagol & Shridhar S. Hiremath : Conceptualisation, draft review and editing; Shridhar S. Hiremath, T. M. Ningaraju & C. N. Lakshminarayana Reddy : Framing research proposal & supervision

Corresponding Author :
C. N. Lakshminarayana Reddy

Received : October 2023
Accepted : December 2023

Abstract

Plant viruses are of significant threat to agricultural productivity worldwide, causing substantial economic losses and impacting food security. Understanding the molecular interactions between plant viruses and their hosts is essential for developing effective strategies for virus control and crop protection. In this study, we employed homology modeling and an *in silico* approach to investigate interactions between chilli host protein, chloroplastic photosystem II protein D1 with the coat protein of cucumber mosaic virus (CMV). CMV coat protein sequences were retrieved from genomic data bank to identify putative host targets. Subsequently, we utilized established protein-protein interaction prediction algorithms and resources to predict potential interactions between CMV proteins and host proteins. This *in silico* study provides a foundation for further experimental investigations, offering potential targets for the development of novel antiviral strategies and crop protection measures. From the obtained docking scores, it was presumed that CMV CP and Photosystem II protein D1 are interacting and established structural-functional relation provided a basis to propose probable host-virus interactions in terms of virus infection and symptom development.

Keywords : Cucumber mosaic virus, Homology modeling, *In silico* analysis, Protein-protein interactions, Antiviral strategies, Crop protection

Plant viral diseases are the major constraints to sustainable and productive agriculture worldwide. These diseases not only reduce the quality of the agricultural produce but also causes yield losses up to 100 per cent (Matthews and Hull, 2002). Among the plant viruses cucumber mosaic virus (CMV) is a severe one distributed in both temperate and tropical regions. It causes substantial damage to a variety of plant species (Kumari *et al*., 2013). CMV is known to infect more than 1200 plant species, including commercially important crops belonging to a diverse taxonomic group, such as chilli, tomato, banana and cucurbits etc. (Carli *et al.,* 2012). CMV belongs to the family, *Bromoviridae,* genus; *Cucumovirus* and is characterized by a tripartite genome consisting of positive-sense, single-stranded RNA genome with one sub-genomic component, RNA4 (Ding *et al*., 1994). It encodes minimum of five proteins: the viral RNA replication proteins 1a and 2a; the RNA-silencing suppressor protein 2b; the cell-to-cell movement protein 3a and the coat protein (CP) (Palukaitis and Garcia Arenal, 2003). This virus is known to be transmitted through various means, which includes, by several species of aphid vectors, mechanical sap, seed and dodder (Abdullahi *et al*., 2001).

Chilli (*Capsicum annuum* L.) is the most important vegetable, spice, medicinal herb and ornamental plant grown across the world. However, CMV has emerged as major threat for its production resulting in low yields (Khan *et al*., 2006). Upon infection, the virus induces various symptoms in chilli plants such as mosaic patterns on leaves, mottling (irregular spots or streaks), yellow discolouration, vein clearing (loss of normal vein colour), leaf deformation, shoestring or leaf narrowing, stunted growth, reduced fruit size and whitish streaks on green fruits. These symptoms collectively contribute to the adverse impact of CMV on chilli plant health and productivity (Vinaykumar *et al*., 2018). The symptoms of CMV infection in different plants are the outcomes of the complex interaction between the host plant and the virus (Stevens, 1983). Many efforts have been made to understand the genomes and genes of mosaic disease causing viruses, however, only a few studies have been attempted to investigate the physiological or molecular basis of host-virus interaction. The chloroplast is a common target of plant viruses and are closely related to the expression of mosaic symptoms (Zhao *et al*., 2016). Indeed, some studies have provided evidence that the coat protein (CP) and the 2b protein of CMV can interact synergistically with other components of CMV to induce viral symptoms (Lewsey *et al*., 2010). This highlights the complexity of the molecular interactions between CMV and its host plants, where multiple viral components may act in concert to elicit specific disease symptoms. In a recent study, Qiu *et al*. (2018) demonstrated that the interaction between chloroplast ferredoxin I protein and the CP of M-CMV was critical for the yellow chlorosis symptom development in tobacco. Determining the specificity of the interaction of chloroplastic protein with the CP of CMV would aid in the construction of symptom development model in chilli and will lead us to understand the physiological and biochemical changes induced by CMV in chilli. The use of bioinformatics in plant pathology has shown tremendous growth in recent years (Hiremath *et al*., 2021).

Therefore, the identification and characterization of proteins involved in these host-virus interactions may underpin the unique and unforeseen roles of these proteins in facilitating host infection. This insight can be pivotal in understanding the mechanisms underlying CMV pathogenicity in different plant species (Kuznetsov and McPherson, 2011).

The present analysis aims to investigate the interactions of the CMV coat protein (CP) with the chloroplastic photosystem II protein D1 in chilli plants by using *in silico* protein-protein docking approaches. This computational approach offers a valuable tool for exploring potential molecular interactions between these proteins, providing insights into the mechanisms underlying CMV infection and its impact on chloroplastic function in chilii plants.

Material and Methods

Selection of host protein

The selection of the host protein for this study was based on previously generated transcriptome and proteome data in CMV-chilli interaction research conducted at the Department of Plant Pathology, CoA, UAS, GKVK, Bengaluru (Vinaykumar, 2020 and Hiremath, 2022). From the generated data, based on functional annotation of a protein, it was observed that the photosystem II protein D exhibited down-regulation in susceptible chilli line (IIHR-2541) when compared to mock treated plants at 24 hours post inoculation of CMV makes it a pertinent candidate for further investigation in this study.

Sequence Retrieval and Computational Analysis of CMV Coat Protein (CP) and Photosystem II Protein D1 and their Physicochemical Properties

The protein sequences of CMV CP (Accession number: WBW02076) and photosystem II protein D1 from chilli (accession number: PHT69894) was retrieved from NCBI Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). To asses the physicochemical properties of these proteins including, amino acid composition, molecular weight, theoretical isoelectric point (pI), total number of positively charged residues, Instability index (II), Aliphatic index (AI) and Grand average of hydropathicity (GRAVY) of a protein the ProtParam program (https://web.expasy.org/protparam/) was utilized (Gasteiger *et al*., 2003).

Prediction of Secondary and Tertiary Structure of CMV CP and Photosystem II Protein D1 of Chilli

Predicting the secondary structure is the base to infer structural properties of an unknown protein, it also helps in predicting the 3D structure. In the case of CMV CP and photosystem II protein D1, the secondary structure was predicted using PSIPRED (http://bioinf.cs.ucl.ac.uk/index.php?id=779) online tool. PSIPRED incorporates two feed-forward neural networks to analyze output data obtained from PSI-BLAST (Position Specific Iterated - BLAST). Further, to derive the three-dimensional structure of CMV CP and Photosystem II protein D1, we conducted homology modeling through the SWISS-MODEL platform (https://swissmodel.expasy.org/) (Water house *et al*., 2018). We selected models based on the criteria such as best Z-score value, Global Model Quality Estimation (GMQE) and QMEANDisCo global statistical parameter. The quality of the predicted structures were analyzed using structural validation algorithms, Structural Analysis and Verification (SAVES 6.0). Additionally, to validate the top-performing 3D model, further, we conducted an evaluation through a Ramachandran plot analysis using the PROCHECK server (Laskowski *et al*., 2006).

***In silico* Docking Analysis of CMV CP and Photosystem II Protein D1 of Chilli**

The CMV CP and photosystem II protein D1 were subjected to *in silico* docking using the program Cluspro server, a widely recognized tool for protein-protein docking, available at https://cluspro.org. Docking models were built by ClusPro 2.06, using default parameters and weighted scores were calculated according to the formula E = 0.40 Erep + -0.40 Eatt + 600 Eelec + 1.00 EDARS for all models and their respective clusters. The best-offered cluster with the most members and the lowest weighted scores docking mode was selected (cluster #0) (Kozakov *et al*., 2017). The predicted models of ClusPro were ranked by cluster size.

Results and Discussion

Characterizing the Physicochemical Properties of Cucumber Mosaic Virus Coat Protein (CMV CP) and Photosystem II Protein D1 of Chilli

Physicochemical properties are considered pivotal in influencing the function and structure of protein sequences. The Expasy ProtParam tool was employed to calculate the physicochemical attributes of the inferred amino acid (aa) sequences of CMV CP and photosystem II protein D1, with the results presented in Table 1. CMV CP is a positively charged protein and photosystem II protein D1 is shown to be negatively charged with calculated molecular weight of the 24.095 and 39.200 KDa, respectively. The calculated isoelectric point (pI) of CMV CP is 9.92, which is greater than 7, indicating a basic nature. In contrast photosystem II protein D1 has pI of 5.12, signifying an acidic nature. The isoelectric point represents the pH at which the protein has no electric charge and this value holds significance in protein purification because of factors such as low solubility, protein stability and compactness (Sahay and Shakya, 2010).

Table 1

Physicochemical properties of CMV CP and photosystem II protein D1 of chilli

Name of the protein	No. of amino acids	Molecular weight (Da)	Theoretical pI	Total no. of negatively charged residues	Total no. of positively charged residues	The instability index (II)	Aliphatic index	Grand average of hydropathicity (GRAVY)
CMV CP	218	24095.52	9.92	21	33	45.88	83.62	-0.353
Photosystem I protein D1	353	39200.88	5.12	28	14	39.38	96.18	0.350

The computed instability index (II) for CMV CP is 45.88 and for photosystem II protein D1 is 39.38. The metabolic stability of a protein within a test tube can be characterized based on the value of II, where a value less than 40 indicating a stable protein while, a value greater than 40 indicates an unstable nature of the protein in the test tube (Gamage *et al*., 2019). Photosystem II protein D1 II value is less than 40, found to be stable *in vitro.* However, the CMV CP was quite unstable and short lived. The aliphatic side chains (alanine, valine, isoleucine and leucine) are responsible for the thermal stability of globular proteins and the aliphatic index (AI) of a protein is defined as the relative volume filled by aliphatic side chains (Panda and Chandra, 2012).

Among the two proteins, photosystem II protein D1 with an AI of 96.18 can be considered to have better thermostability than CMV CP. The grand average hydropathicity (GRAVY) value is calculated by dividing the sum of a protein's hydropathy values by the number of residues in the sequence and it measures the hydrophobicity or hydrophilicity of a protein (Babnigg and Joachimiak, 2010). CMV CP had negative GRAVY rating of -0.353, indicating its hydrophilic nature, whereas, photosystem II protein D1 had positive GRAVY rating of 0.350, suggesting its hydrophobic nature. The computed physico-chemical properties assist in the study of a protein's function and nature and this knowledge is critical during numerous *in vitro* and *in vivo* protein interaction studies (Yu *et al*., 2017).

Prediction of Secondary and Tertiary Structure of CMV CP and Photosystem II Protein D1 of Chilli

The amino acid composition of a protein encodes valuable information regarding its folding and the formation of a stable 3D structure. A protein's 3D structure typically possesses the lowest free energy, which enables, it to interact effectively with other molecules (Deng *et al*., 2018). In our study, we

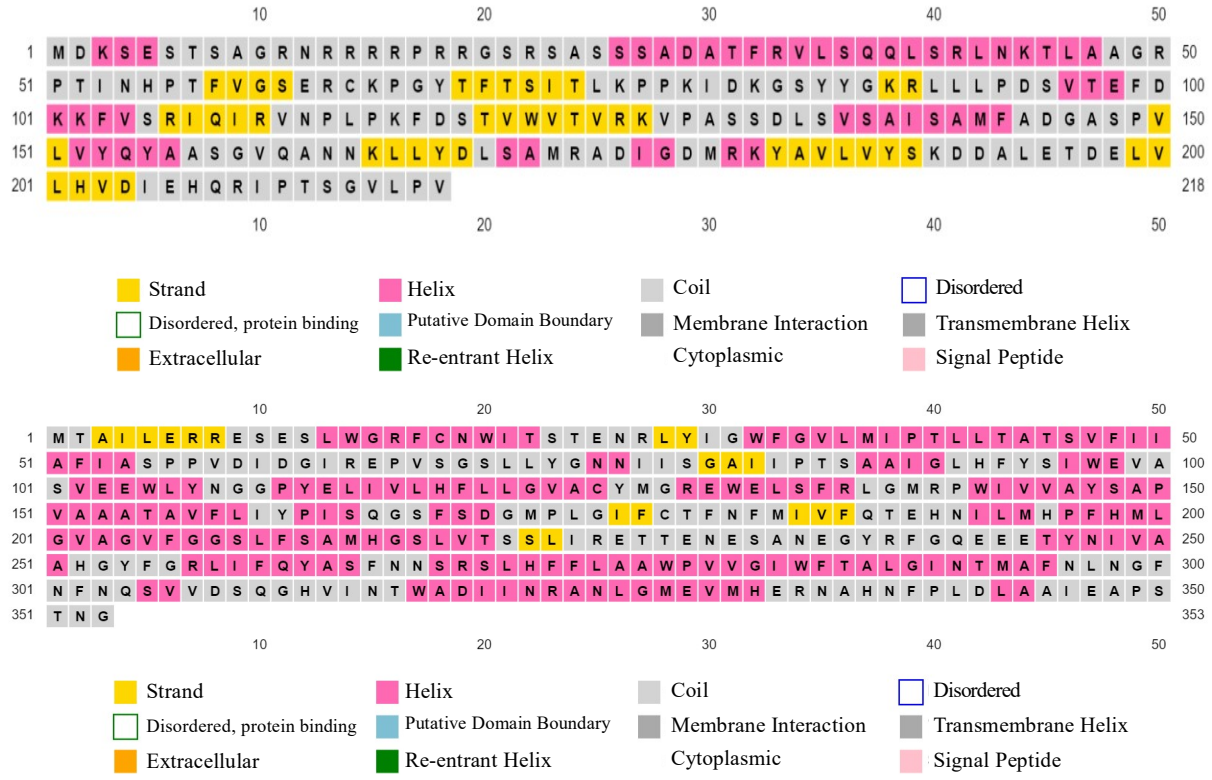

Fig. 1 : Sequence annotation plot for proteins. a. Cucumber mosaic virus coat protein, b. Photosystem II protein D1 of chilli. The different colour code represents different properties of amino acid residues.

The Mysore Journal of Agricultural Sciences

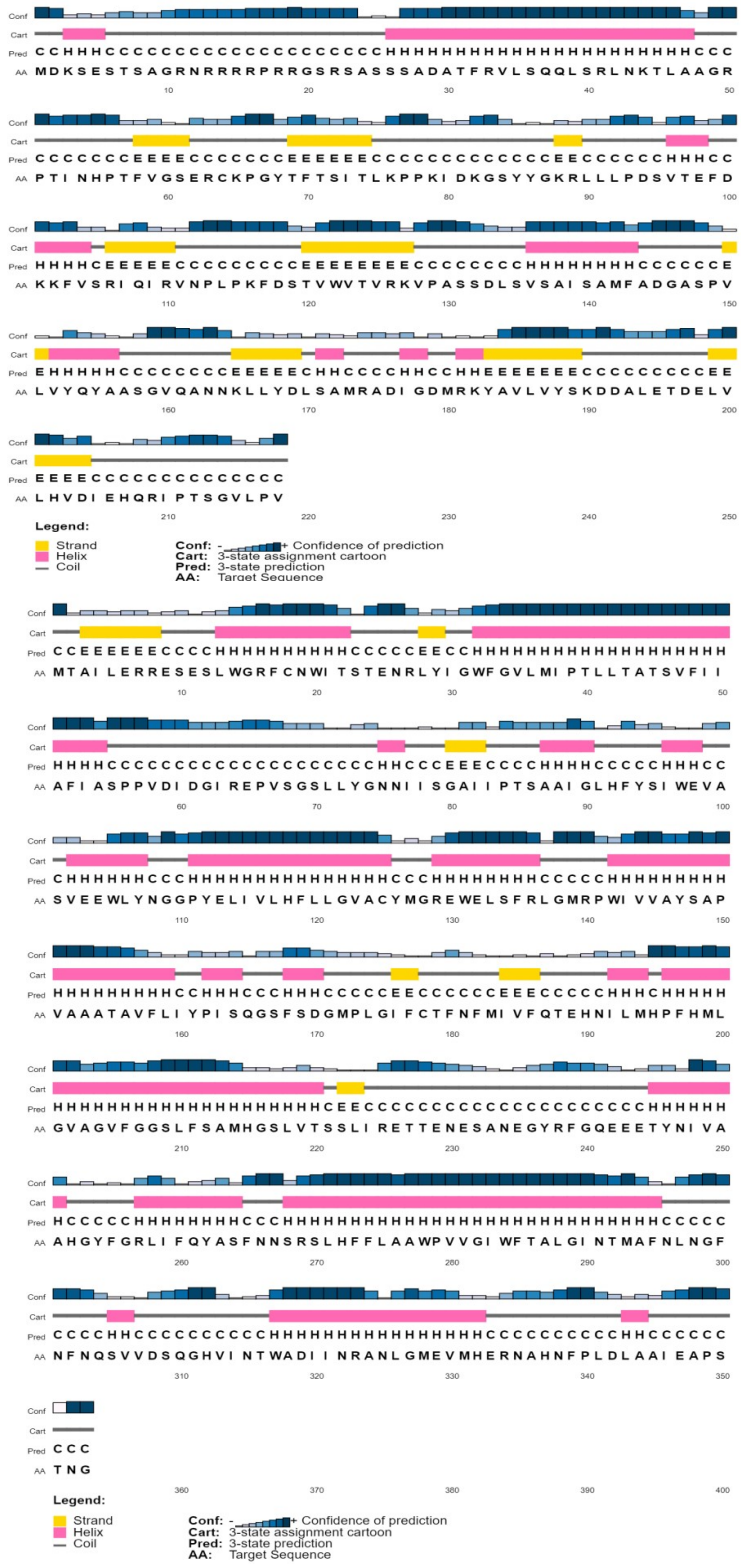

Fig. 2 : PSIPRED cartoon of proteins. a. Cucumber mosaic virus coat protein, b. Photosystem II protein D1 of chilli. The diagrams annotate the query sequence with secondary structure cartoons and confidence value at each position in the alignment. The confidence is given as a series of blue bar graphs

predicted the secondary structure of CMV CP and photosystem II protein D1 of chilli with the aid of PSIPRED online tool. The sequence annotation plot (Fig. 1) illustrates the annotated residues based on the predicted secondary structures.

This graphical representation allows the identification of the disordered residues and binding sites can be obtained. However, the CMV CP and photosystem II protein D1 in the current study was found to have good quality with no disordered residue and binding sites which was also well supported by the PSIPRED cartoon (Fig. 2).

The tertiary structural data for CMV CP and photosystem II protein D1 were unavailable in the PDB database. Consequently, to obtain three-dimensional structure for CMV CP (Fig. 3a,b) and photosystem II protein D1 (Fig. 4a,b), we employed homology modeling by using the SWISS-MODEL platform (https://swissmodel.expasy.org/) (Water house *et al*., 2018). As a template we selected Fny-CMV CP (1f15) and *Arabidopsis thaliana* photo system II protein D1 (7OUI) due to their high sequence similarity to these target proteins.

In the SWISS-MODEL workspace, we generated three models for each protein. Then the model with a Global Model Quality Estimate (GMQE) value of 0.77 and a QMEANDisCo global score of 0.80 ± 0.06 for CMV CP (Fig. 5a), and a model with a GMQE value of 0.84 and a QMEANDisCo global score of 0.77 ± 0.05 for photosystem II protein D1 (Fig. 5b) for subsequent preprocessing were selected. Notably, GMQE and QMEANDisCo global score values ranging from 0 to

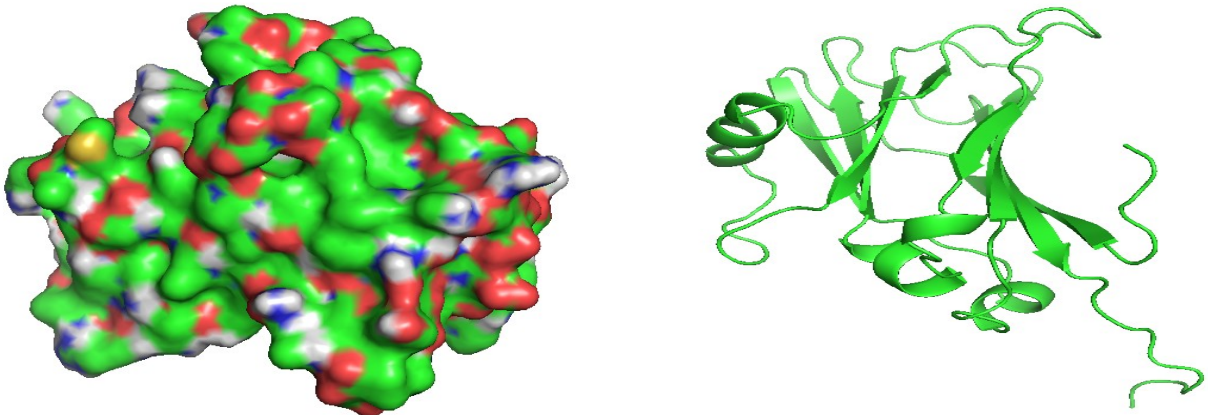

Fig. 3 : The three dimensional visualization of cucumber mosaic virus coat protein deduced using PyMOL software. a. Surface representation, b. Cartoon representation

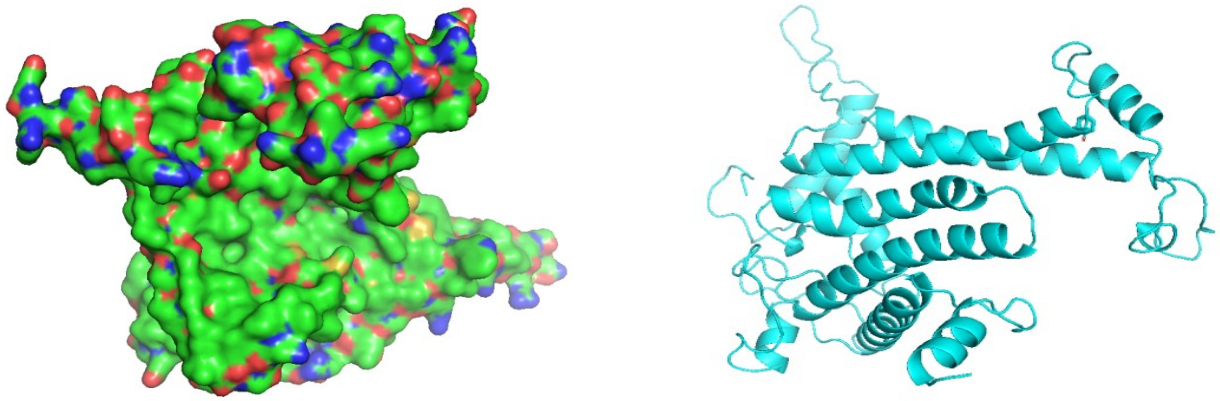

Fig. 4 : The three (3) dimensional visulalization of Photosystem II protein D1 of chillideduced using PyMOL software. a. Surface representation, b. Cartoon representation

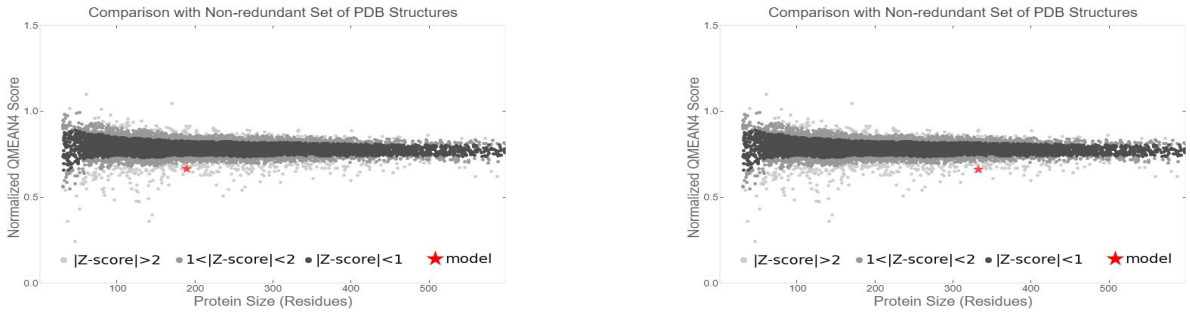

Fig. 5 : Comparison plot : a. Cucumber mosaic virus coat protein, b. Photosystem II protein D1 of chilli. The x-axis shows protein length (number of residues). The y-axis is the 'QMEAN' score. Every dot represents one experimental protein structure. Black dots are experimental structures with a 'QMEAN' score within 1 standard deviation of the mean (|Z-score| between 0 and 1), experimental structures with a |Z-score| between 1 and 2 are grey. Experimental structure that are even further from the mean are light grey. The actual model is represented as a red star

TABLE 2

Evaluating modeled protein structures by ProSA and PROCHECK

	ProSA	PROCHECK regions			
	Z-Score	Most favored %	Additional allowed %	Generously allowed %	Disallowed %
CMV CP	-6.13	87	10.1	2.2	0.7
Photosystem II protein D1	-2.74	91.4	8.6	0	0

1 indicate models of good quality, as established by Waterhouse *et al*. (2018). This reaffirms the high quality of the generated 3D models. The ProSA Z-Scores were determined by comparing the predicted structures against protein structures of the same size obtained by nuclear magnetic resonance and X-ray crystallography. The Z-Scores for the predicted models were found to be -6.13 for CMV CP and -2.74 for photosystem II protein D1 of chilli (Table 2). These Z-scores fall within the range of Z-score typically observed for protein structures of similar size, indicating the reliability and appropriateness of our modeled structures.

The stereochemical quality of a protein model can be assessed by using a Ramachandran plot generated by the PROCHECK program. This algorithm scrutinizes a protein's structure in a plot examining its backbone conformation, by showing the - phi (φ) and psi (ψ) - angles for each residue of a protein. As shown in Table 2, the predicted model of CMV CP exhibited 87 per cent of amino acids in the most favored region, with 10.1 per cent in the additional allowed regions, with 2.2 per cent in the generously allowed regions and 0.7 per cent in the disallowed conformations (Fig. 6a). The predicted model of photosystem II protein D1 of chilli showed presence of 91.4 per cent of amino acids in the most favored region and 8.6 per cent in the additional allowed regions (Fig. 6b). These two models fulfilled all the requirements to be a good quality model.

In silico Docking Analysis of CMV CP and Photosystem II Protein D1 of Chilli

In order to facilitate understanding the nature of interaction between the CMV CP with photosystem II protein D1 of chilli, molecular docking analysis was conducted using the ClusPro server. A visual representation of this 3D interaction is provided in Fig. 7. Among the models generated, the best models with the lowest weighted score were selected, which belonged to a cluster containing 57 members. This selected model exhibited a center energy of -1234.2 and the lowest energy of -1336.6. It's important to note that in molecular docking higher the negative value of interaction is directly proportional to more

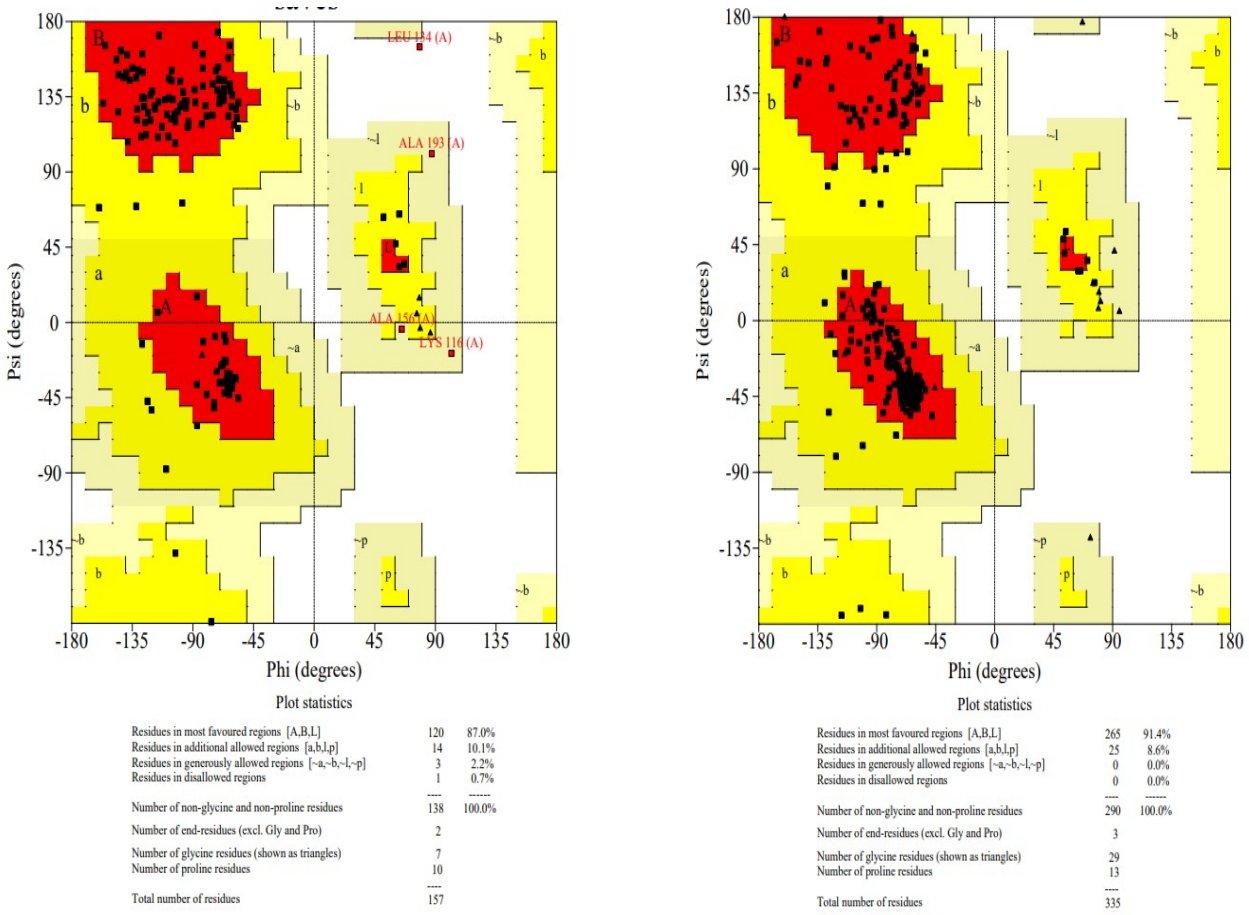

Fig. 6: The Ramachandran plot analysis of proteins. a. Cucumber mosaic virus coat protein, b. Photosystem II Protein D1 of chill. The homology models obtained from SWISS-MODEL were uploaded to PEOCHECK online tool

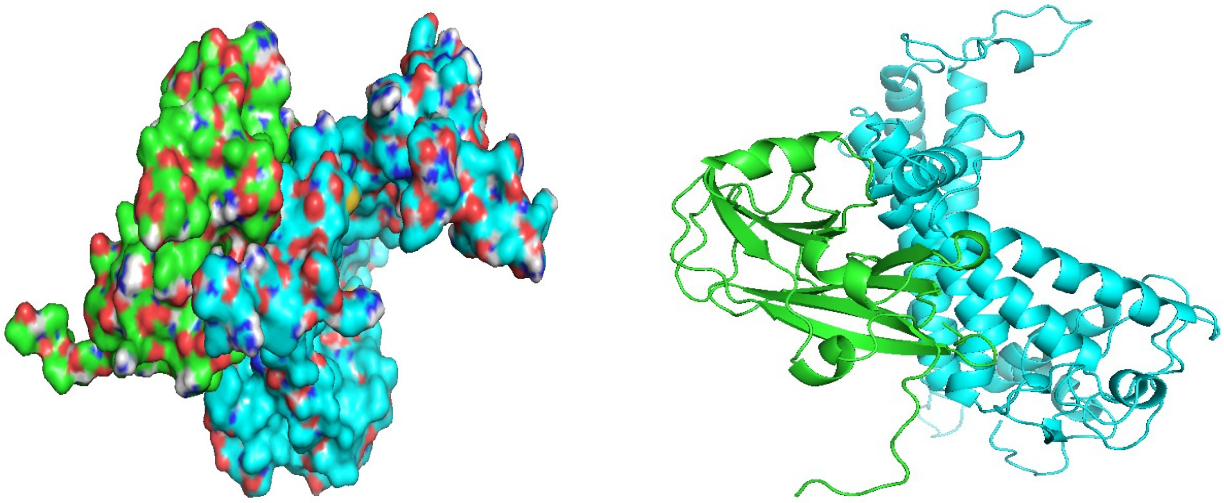

Fig.7: 3D interaction plots of cucumber mosaic virus coat protein with photosystem II protein D1 of chilli determined with the aid of Cluspro server. a. Surface representation, b. Cartoon representation

The Mysore Journal of Agricultural Sciences

stable protein complex (Kamal *et al*., 2019) and obtained high negative values implies possible binding of CMV CP with photosystem II protein D1 of chilli during CMV and chilli interactions.

The interactions between the virus and host plant play a pivotal role in shaping the development of the disease. Our current study suggests that CMV CP interacts with the chloroplastic photosystem II protein D1 of chilli resulting in the impaired function of chloroplast, which eventually can lead to mosaic and chlorosis symptoms on plants. Looking ahead, the future scope of our research aims to understand the molecular basis of these interactions and to get a deep insight into the mechanisms behind these interactions on validating the role of this protein and its interactions using *in vitro* experiments like yeast two-hybrid system which will provide valuable insights into the intricate molecular interactions at play.

References

Abdullahi, I., Ikotun, T., Winter, S., Thottappilly, G. and Atiri, G. I., 2001, Investigation on seed transmission of cucumber mosaic virus in cowpea. *African Crop Sci. J.*, **9**(4) : 677 - 684.

Babnigg, G. and Joachimiak, A., 2010, Predicting protein crystallization propensity from protein sequence. *J. Struct. Funct. Genom*., **11**: 71 - 80.

Carli, M. D., Benvenuto, E. and Donini, M., 2012, Recent insights into plant-virus interactions through proteomic analysis. *J. Proteome Res*., **11**(10) : 4765 - 4780.

Deng, H., Jia, Y. and Zhang, Y., 2018, Protein structure prediction. *Int. J. Mod. Phys. B.,* **32**(18) : 184 - 189.

Ding, S. W., Anderson, B. J., Haase, H. R. and Symons, R. H., 1994, New overlapping gene encoded by the cucumber mosaic virus genome. *Virology*, **198** (2) : 593 - 601.

Gamage, D. G., Gunaratne, A., Periyannan, G. R. and Russell, T. G., 2019, Applicability of instability index for *in vitro* protein stability prediction. *Protein Pept. Lett.,* **26** : 339 - 347.

Gasteiger, A., Gattiker, C., Hoogland, I., Ivanyi, R. D., Appel, A. and Bairoch, 2003, ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*., **31**: 3784 - 3788.

Hiremath, S. S., 2022, Understanding the interaction between cucumber mosaic virus (CMV) and *Capsicum annuum* L. through multi-omics and virus-phytochemicals by computational biology. *Ph.D. Thesis* (Unpub.), Univ. Agric. Sci., Bangalore.

Hiremath, S. S., Vinaykumar, H. D., Koti, P. S., Nandan, M., Shankarappa, K. and Reddy, C. N. L., 2021, *In-silico* investigation of cucumber mosaic virus (CMV) resistance genes and transcription factors involved in CMV hosts generated from chilli-CMV transcriptome data. *Mysore J. Agric. Sci.*, **55**(4) : 355 - 363.

Kamal, H., Fayyaz-Ul-Amir Afsar Minhas, M., Farooq, D. T., Hamza, M., Mustafa, R., Khan, M. Z. and Amin, I., 2019, *In silico* prediction and validations of domains involved in *Gossypium hirsutum* SnRK1 protein interaction with cotton leaf curl multan betasatellite encoded βC1. *Front. Plant Sci.,* **10** : 656 - 670.

Khan, S. M., Shri Krishna Raj, Bano, T. and Vijay Garg, K., 2006, Incidence and management of mosaic and leaf curl diseases in cultivars chilli (*Capsicum annuum* L.). *J. Food Agric. Environ*., **4** : 171 - 174.

Kozakov, D. R., Hall, B., Xia, K. A., Porter, D., Padhorny, C., Yueh, D., Beglov, S. and Vajda, 2017, The ClusPro web server for protein-protein docking. *Nat. Protoc.,* **12** : 255 - 278.

Kumari, R., Bhardwaj, P., Singh, L., Zaidi, A. A. and Hallan, V., 2013, Biological and molecular characterization of cucumber mosaic virus subgroup II isolate causing severe mosaic in cucumber. *Indian J. Virol.*, **24** : 27 - 34.

Kuznetsov, Y. G. and Mcpherson, A., 2011, Atomic force microscopy in imaging of viruses and virus-infected cells. *Microbiol Mol. Biol. Rev.,* **75** : 268 - 285.

Laskowski, R. A., Macarthur, M. W. and Thornton, J. M., 2006, PROCHECK : Validation of protein-structure coordinates.

LEWSEY, M. G., GONZALEZ, I., KALININA, N. O., PALUKAITIS, P., CANTO, T. AND CARR, J. P., 2010, Symptom induction and RNA silencing suppression by the cucumber mosaic virus 2b protein. *Plant Signal. Behav.,* **5** (6) : 705 - 708.

MATTHEWS, R. E. F. AND HULL, R., 2002, *Matthews' plant virology*. Gulf professional publishing.

PALUKAITIS, P. AND GARCIA-ARENAL, F., 2003, Cucumoviruses. *Adv. Virus Res*., **62** : 241 - 323.

PANDA, S. AND CHANDRA, G., 2012, Physicochemical characterization and functional analysis of some snake venom toxin proteins and related non-toxin proteins of other chordates. *Bioinformation*, **8** : 891 - 896.

QIU, Y., ZHANG, Y., WANG, C., LEI, R., WU, Y., LI, X. AND ZHU, S., 2018, Cucumber mosaic virus coat protein induces the development of chlorotic symptoms through interacting with the chloroplast ferredoxin I protein. *Sci. Rep*., **8** (1) : 1205 - 1216.

SAHAY, A. AND SHAKYA, M., 2010, *In silico* analysis and homology modelling of antioxidant proteins of spinach. *J. Proteomics Bioinform*., **3** : 148 - 154.

STEVENS, W. A., 1983, Virology of Flowering Plants. *Springer US*.

VINAYKUMAR, H. D., 2020, Characterization of cucumber mosaic virus (CMV) and understanding the molecular mechanisms of CMV and chilli interaction through transcriptomics. *Ph.D. Thesis*, Univ. Agric. Sci., Bangalore.

VINAYKUMAR, H. D., REDDY, C. N. L. AND KRISHNAREDDY, M., 2018, Molecular detection of cucumber mosaic virus infecting chilli (*Capsicum annuum* L). *Mysore J. Agric. Sci*., **54** (4) : 738 - 743.

WATERHOUSE, A., BERTONI, M., BIENERT, S., STUDER, G., TAURIELLO, G., GUMIENNY, R., HEER, F.T., DE BEER, T.A.P., REMPFER, C., BORDOLI, L., 2018, SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acid Res*., **46** : 296 - 303.

YU, L., ZHANG, Y., GUTMAN, I., SHI, Y. AND DEHMER, M., 2017, Protein sequence comparison based on physicochemical properties and the position-feature energy matrix. *Sci. Rep*., **7** : 46237 - 46237.

ZHAO, J., ZHANG, X., HONG, Y. AND LIU, Y., 2016, Chloroplast in plant-virus interaction. *Front Microbiol.,* **7** : 1565 - 1575.