International Journal of Zoology and Applied Biosciences **ISSN: 2455-9571**
Volume 9, Issue 4, pp: 48-51, 2024
https://doi.org/10.55126/ijzab.2024.v09.i04.008 **http://www.ijzab.com**

Research Article

CHARACTERIZATION OF CAPSID PROTEIN OF PORCINE CIRCOVIRUS 4 BY IN-SILICO APPROACH

[1]Arpita Bharali, [1]Nagendra Nath Barman, [1]Jayashree Sarma, [1]Durlav Prasad Bora, and [2*]Lukumoni Buragohain

[1]Department of Veterinary Microbiology, College of Veterinary Science, Assam Agricultural University, Khanapara, Guwahati, Assam
[2]Department of Animal Biotechnology, College of Veterinary Science, Assam Agricultural University, Khanapara, Guwahati, Assam

Article History: Received 22nd April 2024; Accepted 11th June 2024; Published 31st July 2024

ABSTRACT

Porcine Circovirus 4 (PCV4) has been detected in pigs from China only in 2019. PCV4 has been found to be associated with several forms of clinical syndromes in pigs. Like other Porcine Circoviruses, the only structural protein of PCV4 is the capsid protein. Therefore, it could be a key vaccine candidate or diagnostic marker. The present study aims to characterize the capsid protein of PCV4 by using several bioinformatics tools. The in-silico analysis revealed that the molecular weight of PCV4 capsid protein is about 27.3 kDa which is made up of 228 amino acids. The capsid protein of PCV4 contains a nuclear localization signal at N-terminal and several post-translational modifications could be predicted at different residues of the protein. The secondary structure of capsid protein contains coil, beta strand, and alpha helix. The 3D structure of the protein was predicted with a high level of confidence. The capsid protein of PCV4 also contains several potential B-cell epitopes. Thus, the in-silico characterization of the capsid protein will be beneficial in designing diagnostics and vaccine candidates in future to combat PCV4 infection in pigs.

Keywords: Bioinformatics analysis, Capsid protein, PCV4, Pigs.

INTRODUCTION

Piggery sector has tremendous scope in India, particularly in north-eastern region. Pig farming provides livelihood to several families. Pig meat, 'pork' is a good source of many nutrients and it is becoming one of the popular meats among the new-generation youths. However, pigs are infected with several viruses and one of the viruses is porcine circovirus which cause huge economic loss to the pig industry worldwide. Till date, four types of porcine circoviruses (PCVs) have been identified which are PCV1, PCV2, PCV3 and PCV4. Porcine circovirus type 4 (PCV4) is the newest type among the porcine circoviruses and it has been detected in several clinical cases and suspected to be associated with several clinical syndromes (Wang *et al.*, 2022). PCV4 was first detected from infected pigs in 2019 from Hunan province of China (Zhang *et al.*, 2020). The PCV4 is a non-enveloped, single-stranded DNA virus belonging to the genus *Circovirus* and family *Circoviridae* (Zhang *et al.*, 2020).The length of PCV4 genome is about 1700 nucleotides. Like other PCVs, the genome of PCV4 also contains two major genes, *rep* and *cap* which encode replicas and capsid protein (Wang *et al.*, 2021). Capsid protein is the only structural protein of PCV4 and it is the major immunogenic protein. Hence, capsid protein is a promising target for development of vaccine candidate or diagnostic marker against PCV4. This study aims to characterized the structure of capsid protein by using different bioinformatics tools so that the further knowledge related to the protein could be revealed.

MATERIALS AND METHODS

Sequence retrieval and consensus sequence preparation

A total of ninety-five (95) complete coding sequences (CDSs) of *cap* gene (ORF 2) belonging to different strains

*Corresponding Author: Dr. Lukumoni Buragohain, Department of Animal Biotechnology, College of Veterinary Science, Assam Agricultural University, Khanapara, Guwahati, Assam. Email: lukumoni@gmail.com.

of PCV4 were retrieved in fasta format from NCBI virus database (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/). A consensus sequence of *cap* gene was prepared from 95 sequences by employing online tool EMBOSS Cons (https://www.ebi.ac.uk/jdispatcher/msa/emboss_cons).

Physico-chemical properties of capsid protein

The consensus *cap* gene sequence was translated into amino-acid sequence by using Translate tool available in Exapsy server (https://web.expasy.org/translate/). ProtParam tool of Expasy (http://web.expasy.org/protparam/) was used to compute different physico-chemical properties of PCV4 capsid protein such as molecular weight, theoretical isoelectric point (pI), numbers of positively and negatively charged amino acids, grand average hydropathy (GRAVY), aliphatic index, and instability index.

Prediction of Secondary structure, domain and modifications of PCV4 capsid protein

The secondary structure of PCV4 capsid protein was predicted using PSIPRED server (Buchan and Jones, 2019). Secretory signal peptide(s) and cleavage site(s) present in the capsid protein was predicted using the Signal P 5.0 server (https://services.healthtech.dtu.dk/services/SignalP-5.0/). Presence of any transmembrane domain was predicted by using Deep TMHMM online software (https://dtu.biolib.com/DeepTMHMM). NLStradamus program (Nguyen Ba *et al.*, 2009) was used to identify the probable nuclear localization signal embedded in capsid protein of PCV4. The probable post-translational modification sites such as O-linked glycosylation, N-linked glycosylation, phosphorylation, glycation, N-terminal acetylation, and C-mannosylation present in the capsid protein of PCV4 was predicted by using online servers Net OGlyc 4.0 (Steentoft *et al.*, 2013), NetNGlyc1.0 (Gupta and Brunak, 2002), Net Phos 3.1(Blom *et al* ., 1999), Net Glycate 1.0 (Johansen *et al.*, 2006), Net Acet 1.0 (Kiemer *et al.*, 2005), and NetCGlyc 1.0 (Julenius, 2007), respectively.

Modelling of 3D structure of PCV4 capsid protein

Protein Homology/analog Y Recognition Engine V 2.0 (Phyre2) web portal (Kelley *et al.*, 2015) was deployed to predict the 3-Dimensional (3D) structure of capsid protein of PCV4. The older version Phyre was replaced by Phyre2 which uses advanced algorithm to predict 3D structure, function, and ligand binding sites in a protein along with amino-acid mutation effect in protein sequence.

Prediction of B-cell epitopes in capsid protein of PCV4

Linear B-cell epitope (s) present in the capsid protein were predicted by using Bepipred Linear Epitope Prediction 2.0 program (Jespersen *et al.*, 2017) based on consensus amino acids sequence of the capsid protein.

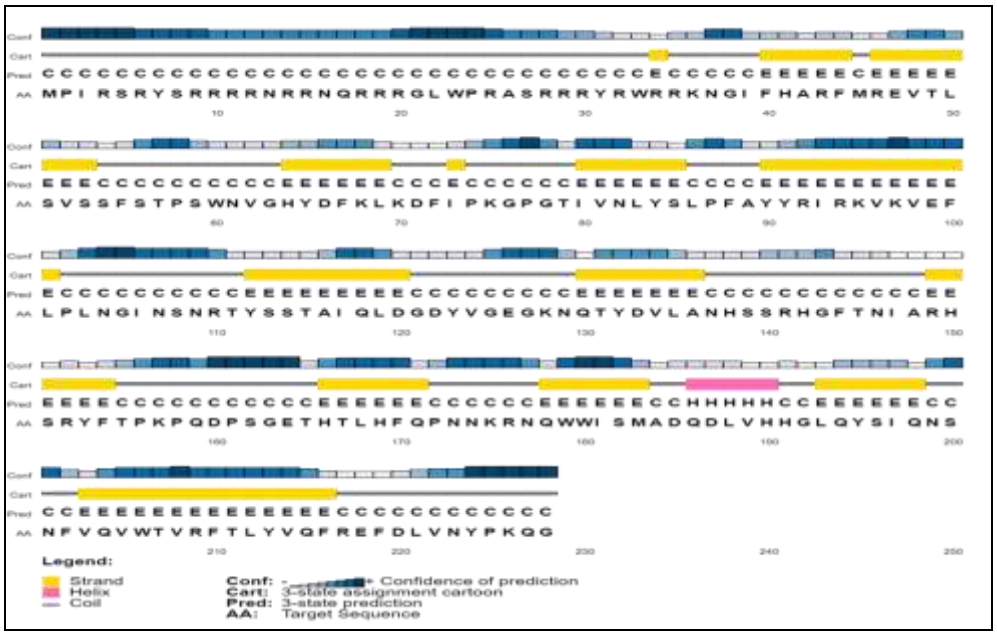

Figure 1. Secondary structure of PCV4 capsid protein predicted by PSIPRED program. The solid grey horizontal line indicates random coil, yellow bars represents beta strands and pink bars represents the alpha helixes. The individual coloured bars in gradient (white to blue) represent confidence of prediction of each amino acid.

RESULTS AND DISCUSSION

The consensus *cap* gene sequence generated from 95 CDS of PCV4 strains available in public database consist of 687 nucleotides and upon translation it encodes capsid protein with 228 amino acids. The approximate molecular weight of PCV4 capsid protein was estimated to be 27.3 kDa with theoretical pI of 11.05. The predicted molecular weight is almost equivalent to weight of the original protein, as Wang *et al.* (2021) documented the molecular weight of PCV4 capsid protein to be ~28 kDa which was expressed in *E. coli*. The number of positively and negatively charged amino acids present in capsid protein were 39 (Arg + Lys) and 14 (Asp + Glu), respectively. The grand average hydropathy (GRAVY), aliphatic index, and instability index were found to be -0.897, 61.93, and 52.96, respectively. The capsid protein of PCV4 is devoid of secretory signal peptide and transmembrane domain. Like other PCVs, the capsid protein of PCV4 also has a nuclear localization signal at N-terminal end (Zhou *et al.*, 2021) which spans from amino acid positions 6 to 37 ($_{6}$RYSRRRRNRRNQRRRGLWPRASRRRYRWRRKN$_{37}$) as predicted by NLStradamus program. However, the important NLS motif of PCV4 capsid protein is $_{8}$RRRR-RR-RRR$_{20}$ (Zhou *et al.*, 2021). The dominant secondary structure of PCV4 capsid protein was observed to be random coil followed by beta strand and a minimum portion was observed to be alpha helix (Figure 1). The different types of post-translational modifications that were predicted in capsid protein of PCV4 by deploying different online programs are highlighted in table 1.

The 3D structure of PCV4 capsid protein was successfully generated by deploying Phyre2 web portal. The best 3D model of the protein (Figure 2) was predicted with 100% confidence and showed 84% query coverage (192 residues) and 49% identity with bat circovirus capsid protein (PDB ID:6RPK). The 3D structure of PCV4 capsid protein predicted in this study also similar to jelly roll structure as demonstrated by Wang *et al*. (2021) Seven (7) numbers of potential linear B-cell epitopes were predicted in the capsid protein of PCV4 by Bepipred Linear Epitope Prediction 2.0 program (Figure 3), and similarly, Wang *et al*. (2021) had identified 5 numbers of linear B-cell epitopes which were also found in this study. Moreover, the linear epitope that is present between position 5 and 35 (Wang *et al*., 2022) is also being predicted in this study. Thus, the seven linear B-cell epitopes predicted in the capsid protein of PCV4 could be critical and promising diagnostic markers as well as vaccine candidates. However, the immunogenicity of these epitopes needs to be validated by other immunoassays.

Table 1. List of in-silico predicted post-translational modification of PCV4 capsid protein.

Sl No.	Modifications	Amino acid Positions	Program
1	O-linked glycosylation	5, 8, 151, 155	NetOGlyc 4.0
2	N-linked glycosylation	109	NetNGlyc 1.0
3	Phosphorylation	8, 27, 49, 51, 54, 57, 112, 113, 114, 123, 132, 139, 145, 151, 155, 162, 182, 207, 211,213	NetPhos 3.1
4	Glycation	36, 68, 128	NetGlycate 1.0
5	N-terminal Acetylation	Nil	NetAcet 1.0
6	C-mannosylation	Nil	NetCGlyc 1.0

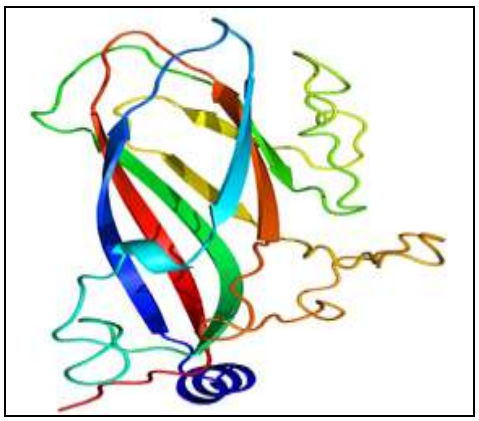

Figure 2. 3D structure of PCV4 capsid protein generated by Phyre2 web portal.

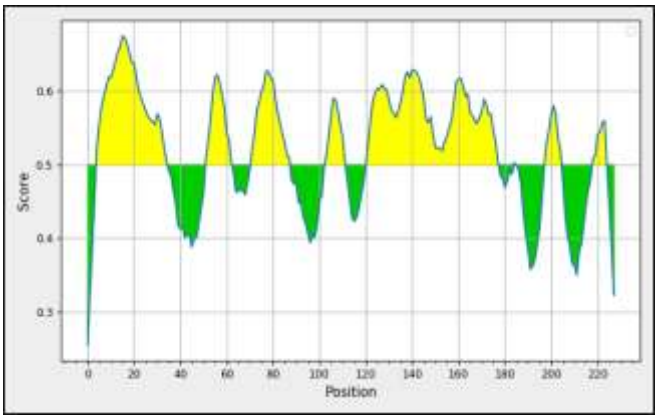

Figure 3. Linear B-cell epitopes predicted in Capsid protein of PCV4 by Bepipred Linear Epitope Prediction 2.0 program.

The x-axis represents the amino acid position and y-axis represents the score. The region in yellow above 0.5 score predicted to form potential B-cell epitopes.

ACKNOWLEDGEMENT

The authors are grateful to the Director of Research (Veterinary), Assam Agricultural University, Khanapara, Guwahati, Assam for providing the facilities to carry out the research.

REFERENCES

Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology*, 294, 1351–1362.

Buchan, D. W. A., & Jones, D. T. (2019). The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research*, 47, W402–W407.

Gupta, R., &Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symposium on Biocomputing. *Pacific Symposium on Biocomputing*, 310–322.

Jespersen, M. C., Peters, B., Nielsen, M., &Marcatili, P. (2017). BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research*, 45, W24–W29.

Johansen, M. B., Kiemer, L., &Brunak, S. (2006). Analysis and prediction of mammalian protein glycation. *Glycobiology*, 16, 844–853.

Julenius K. (2007). NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. *Glycobiology*, 17, 868–876.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10, 845–858.

Kiemer, L., Bendtsen, J. D., &Blom, N. (2005). NetAcet: prediction of N-terminal acetylation sites. *Bioinformatics*, 21, 1269–1270.

Nguyen Ba, A. N., Pogoutse, A., Provart, N., & Moses, A. M. (2009). NL Stradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics*, 10, 202.

Steentoft, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T., Lavrsen, K., Dabelsteen, S., Pedersen, N. B., Marcos-Silva, L., Gupta, R., Bennett, E. P., Mandel, U., Brunak, S., Wandall, H. H., Levery, S. B., & Clausen, H. (2013). Precision mapping of the human O-GalNAcglycoproteome through Simple Cell technology. *The EMBO Journal*, 32, 1478–1488.

Wang, D., Mai, J., Lei, B., Zhang, Y., Yang, Y., & Wang, N. (2021). Structure, Antigenic Properties, and Highly Efficient Assembly of PCV4 Capsid Protein. *Frontiers in Veterinary Science*, 8, 695466.

Wang, D., Mai, J., Yang, Y., Xiao, C. T., & Wang, N. (2022). Current knowledge on epidemiology and evolution of novel porcine circovirus 4. *Veterinary Research*, 53, 38.

Zhang, H. H., Hu, W. Q., Li, J. Y., Liu, T. N., Zhou, J. Y., Opriessnig, T., & Xiao, C. T. (2020). Novel circovirus species identified in farmed pigs designated as Porcine circovirus 4, Hunan province, China. *Transboundary and Emerging Diseases*, 67, 1057–1061.

Zhou, J., Qiu, Y., Zhu, N., Zhou, L., Dai, B., Feng, X., Hou, L., & Liu, J. (2021). The Nucleolar Localization Signal of Porcine Circovirus Type 4 Capsid Protein Is Essential for Interaction With Serine-48 Residue of Nucleolar Phosphoprotein Nucleophosmin-1. *Frontiers in Microbiology*, 12, 751382.