Journal of Advanced Bioinformatics Applications and Research
ISSN 0976-2604.Online ISSN 2278 – 6007
Vol 4, Issue 1, 2013, pp1-5
http://www.bipublication.com

INSILICO ANALYSIS AND 3D STRUCTURE PREDICTION OF MATRIX PROTEIN OF NIPAH VIRUS

Sachin Gupta [2], Santosh Kumar[1*], Sarika Saxena [2] ,Rajkumar Yadav[2], Ankur Thakur[2], O. P. Verma[3] and Chandrabhan Seniya[2]

[1]School of Life Sciences, ITM University, Gwalior, M.P,India
[2] Department of Biotechnology, Madhav Institute of Technology and Science Gwalior-05, M.P,India
[3]Jacob School of Biotechnology & Bioengineering, Sam Hogginbottom Institute of Agricultur, Technology & Science, Allahabad, U.P., India
[*]**Corresponding author:** Email: santosh_mtech@rediffmail.com, Tel: +091-0751-2440060 ; Fax: +091-0751-2440058

[Received-15/01/2013, Accepted-11/02/2013]

ABSTRACT:

Nipah virus (NiV) is a newly emerging paramyxovirus that causes lethal infections in various species including humans. The case fatality rate of Nipah virus encephalitis was found to be 70% in India and Bangladesh. NiV have single-stranded negative-sense RNA that encodes six viral proteins out of which Matrix protein of Nipah virus plays a vital role in virus assembly and budding. Physiochemical characterization was done to interpret properties like pI, EC, AI, GRAVY and instability index. The 3D structure of this protein is not available yet. So homology modeling was performed to generate good quality models. The assessment of generated three dimensional structure against structure verification tools PROCHECK and WHATIF showed that model generated by Swiss Model was more acceptable to that by GENO 3D. The predicted model can be used in structure based drug designing and vaccine development.

Keywords: *Nipah Virus (NiV)*

[I] NTRODUCTION

Nipah virus (NiV) is a newly emerging paramyxovirus that has recently appeared from fruit bats of the genus *Pteropus* to cause fatal diseases in humans [8, 9, 5]. With their exceptional wide host range, their zoonotic potential and their ability to cause fatal diseases in animals and humans, henipaviruses differ from all other known paramyxoviruses and are classified as Biosafety Level 4 (BSL4) pathogens [8]. The disease appeared in pigs as acute respiratory distress syndrome[12,21]. Infected human, presented with fever, headache and drowsiness that could develop to fatal central nervous system infection in about 40% of the patients [6, 14]. Recent Hendra virus outbreak in Queensland, Australia (Aug-Sep 2009) have killed 3 horses and one veterinarian, and affected horse farms and potentially infected individuals were quarantined.

It was identified as the causative agent responsible for an outbreak of severe encephalitis in Malaysia and Singapore that began in 1998 and continued into 1999 having a case-fatality rate of 40% [5]. Nipah virus encephalitis had Near about 70% case fatality rate in India in 2001 and Bangladesh from 2001 to 2004, and spread of the virus from human to human was suspected [15,19].

NiV is a negative-stranded, ssRNA that comprise a new genus Henipavirus within the family *Paramyxoviridae.* This is a family of viruses with lipid envelopes derived from the host cell membrane. The genome contains six principle genes: nucleocapsid (N), phosphoprotein (P), polymerase (L), matrix (M), fusion (F) and attachment (HN, H or G) proteins [10]. Paramyxoviruses replicate in the cytoplasm, and progeny virions are released from the plasma membrane of the host cell. Viral assembly and budding are directed by the matrix protein (M), a major structural protein underlying the viral envelope [10, 19,24].

Due to availability of sequential and structural information to the researchers role of bioinformatics in studying different biotechnological problems is increasing day by day. A huge number of computational tools available makes *In silico* approaches less time consuming and money saving than that of experimental methods. in this paper the main aim of the study is to do *Insilico* analysis and 3D structure prediction of matrix protein of Nipah virus. The structure for this protein has not been reported yet. To understand its various structural features a good quality 3dimensional model was constructed.

[II] MATERIALS AND METHODS

2.1. Retrieval of target sequence

The amino acid sequence of the matrix protein of Nipah virus was obtained from the sequence database of NCBI (http://www.ncbi.nlm.nih.gov/protein/NP_112025.1).It was ascertained that the three-dimensional structure of the protein was not available in Protein Data Bank [4]. Hence the present work for developing the 3D model of the matrix protein from *Nipah virus* was undertaken. The protein is 352 amino acids in length.

2.2.Physico-chemical characterization

The values of theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient[13], instability index[15], aliphatic index [17] and grand average hydropathy (GRAVY) [18]were computed. For physico-chemical characterization, using the Expasy's ProtParam server [11] (http://us.expasy.org/tools/protparam.html). The results were shown in Table 1.

2.3.Secondary structure prediction

Secondary structure has been predicted using PHYRE 2 (http://www.sbg.bio.ic.ac.uk/phyre2) software where the FASTA format of the sequence was given as input. SOPMA [12] was employed for calculating the secondary structural features of the protein sequence considered for this study. The results were presented in Table 2.

2.4.Model building and quality assessment

The modeling of the three dimensional structure of the protein was done using two homology modeling programs, Geno 3D [7] and SwissModel [3] The overall stereochemical property of the protein was assessed by Ramchandran plot analysis[23]. The evaluation of structure models obtained from the two software tools was performed by using PROCHECK [20]. The models were further checked with WHAT IF [25] for Standard bond lengths and bond angles determination. The results of PROCHECK were shown in Table 3.

[III] RESULTS AND DISCUSSION

In present study, the protein sequence of matrix protein of Nipah virus was retrieved from NCBI Entrez sequence search in FASTA format and used as query sequence for homology modeling.Physiochemical Parameters computed using Expasy's ProtParam tool was represented in Table 1. A protein having instability index smaller

than 40 is predicted as stable, on the other hand a value above 40 predicts that the protein may be unstable[15].Instability index of 29.53 indicates the stable nature of protein. The high extinction coefficient (43235) indicates presence of high concentration of Cys, Trp and Tyr. The aliphatic index is considered as a positive factor for the increase of thermal stability. High aliphatic index (90.26) of query protein suggests that the protein may be stable for a wide temperature range. The Grand Average hydropathy (GRAVY) value is low (-0.211) and indicates the possibility of better interaction with water.

S.NO.	Property	Value
1.	Number of amino acides	352
2.	Molecular weight	39928.2
3.	Theoretical pl	9.31
4.	Total number of negatively charged residues(Asp+Glu)	36
5.	Total number of positively residues (Arg+Lys)	48
6.	Extinction coefficient	43235
7.	Extinction coefficient*	42860
8.	Instability index	29.53
9.	Aliphatic index	90.26
10.	Grand average of hdropathicity	-0.211

Table 1-Parameters computed using Expasy's ProtParam tool

The secondary structure of Nipah virus matrix protein was predicted by two software namely SOPMA (Self Optimized Prediction Method with Alignment) and phyre 2. SOPMA predicts 69.5% of amino acids correctly to describe secondary structure prediction[12].the results of SOPMA are presented in Table 2.These results show higher number of random coils in comparison to other secondary structure elements (alpha helix, extended strand and beta turns).default parameters (Window width: 17, similarity threshold: 8 and number of states: 4) were taken by SOPMA for secondary structure prediction. Secondary structure and disorder prediction was made using phyre 2 which is shown in figure 1.

S.NO.	parameters	Value(%)
1.	**Alpha helix**	20.17
2.	3_{10}helix	0.00
3.	Pi helix	0.00
4.	Beta bridge	0.00
5.	Extended strand	23.58
6.	Beta turn	5.68
7.	Bend region	0.00
8.	Random coil	50.57
9.	Ambiguous state	0.00
10.	Other state	0.00

Table 2 Calculated secondary structure elements by SOPMA.

Three dimensional structures of proteins are predicted due to unavailability of such data. There is no experimental structure found for the protein considered. The homology modelling of the protein was done using two programs Geno3d and Swiss Model. The results obtained from these programs were compared in table 3. This comparison suggests that the model generated by Swiss model is much better than the model generated by Geno 3d. Finally chosen model was visualized by Rasmol (figure 3).

The evaluation of predicted structure generated by Swiss Model for the stereochemical quality was done using Ramchandran map calculations done with the PROCHECK (figure 2). The 88.3% of residues were found in the core right handed alpha helices(A),beta sheets (B)and left-handed alpha helix(L) region.10.7% of residues were found in the allowed right- handed alpha helix(a),beta sheets(b) and left -handed alpha helices regions. The 0.7% of the residues were found in the generously allowed alpha helices (~a), beta sheets (~b), left handed alpha helices (~l) and epsilon (~p) regions. The 0.3% of the residues was found to be localized at the disallowed regions. Above results indicate to a good quality of predicted model.

The analysis of modeled structure of matrix protein of Nipah virus by structure verification server WHATIF revealed such RMS Z-score (0.965) which was nearly meeting the required score for a high quality model.

Server	Parameters	Value (%)
Geno 3d	Residues in the most Favoured Region	**60.4**
	Residues in additionally allowed region	**29.5**
	Residues in generously allowed region	**6.1**
	Residues in disallowed region	**4.0**
Swiss model	Residues in the most Favoured Region	**88.3**
	Residues in additionally allowed region	**10.7**
	Residues in generously allowed region	**0.7**
	Residues in disallowed region	**0.3**

Table 3 Ramachandran plot calculation and Comparative analysis of the models from Geno3D and Swiss-model computed with the PROCHECK program

Figure 1 secondary structure and disorder prediction done with PHYRE 2.

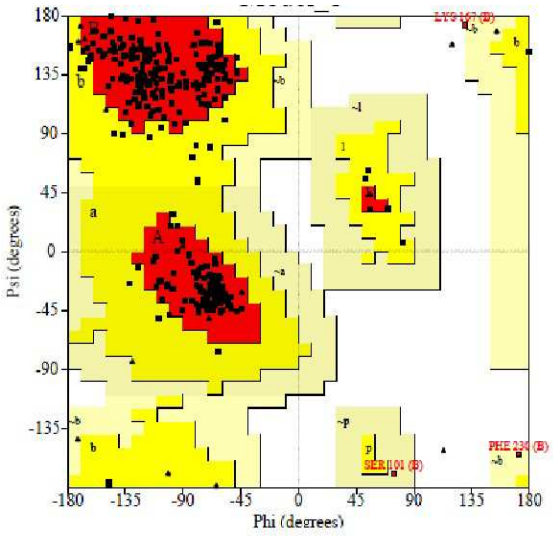

Figure 2: Ramachandran plot of Matrix protein of Nipah virus generated using Procheck software

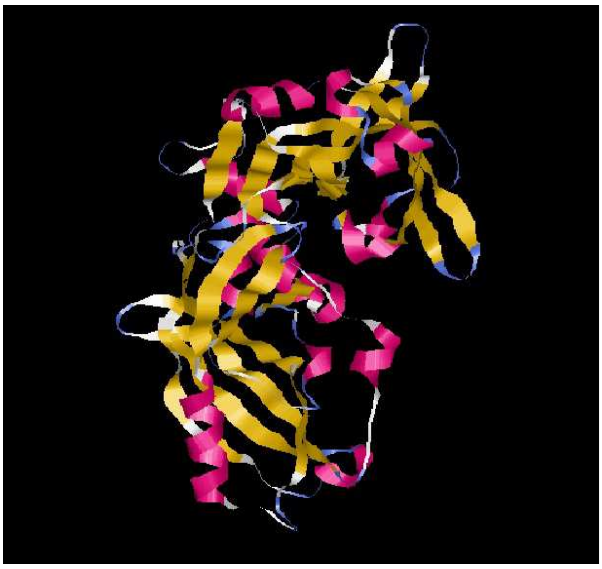

Figure 3: Modeled Structure of matrix protein of Nipah virus

[V] CONCLUSION

On the basis of various structural and physiochemical parameters assessment, it can be concluded that the predicted three dimensional structure of matrix protein of Nipah virus is stable. Since no effective therapeutic or vaccine is available for Nipah virus encephalitis. Structural information of this model can be effectively used and can be further implemented in future drug designing

ACKNOWLEDGEMENT

We would like to express our heartfelt feeling of gratefulness to all those who directly or indirectly helped us in this work.

REFERENCES

1. Anonymous. Outbreak of Hendra-like virus – Malaysia and Singapore,[1998-1999]. *MMWR Morb Mortal Wkly Rep* 1999; 48: 265-9.
2. Anonymous. Update: Outbreak of Nipah virus - Malaysia and Singapore, [1999]. *MMWR Morb Mortal Wkly Rep* 1999; 48:335-7.
3. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22: 195-201

4. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN and Bourne PE, The Protein Data Bank. Nucleic Acids Res, 28 (1): 235-42, 2000.
5. Chua KB, Bellini WJ, Rota PA, Harcourt BH, Tamin A, et al. [2000] Nipah virus: a recently emergent deadly paramyxovirus. Science. ; 288:1432–1435.
6. Chua KB, Goh KJ, Wong KT, *et al*.[1999] Fatal encephalitis due to Nipah virus among pig-farmers in Malaysia. *Lancet*; 354:1257-9.
7. Combet C, Jambon M, Deleage G, Geourjon C (2002) Geno3D: Automatic comparative molecular modelling of protein. Bioinformatics 18: 213-214.
8. Eaton BT, Broder CC, Middleton D, Wang LF. [2006] Hendra and Nipah viruses: different and dangerous.Nat Rev Microbiol.;4:23–35.
9. Field H, Young P, Yob JM, Mills J, et al. [2001] The natural history of Hendra and Nipah viruses.Microbes Infect. ;3:307–314.
10. Garoff H, Hewson R, Opstelten DJ[1998]. Virus maturation by budding. Microbiol Mol Biol Rev.;62:1171–1190.
11. Gasteiger E (2005) Protein Identifi cation and Analysis Tools on the ExPASy Server. In: John M. Walker, The Proteomics Protocols Handbook, Humana Press 571-607.
12. Geourjon C, Deléage G (1995) SOPMA: signifi cant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11: 681-684.
13. Gill SC, Von Hippel PH (1989) Extinction coeffi cient. Anal Biochem 182: 319-328
14. Goh KJ, Tan CT, Chew NK, *et al* [2000]. Clinical features of Nipah virus encephalitis among pig farmers in Malaysia. *N Engl J Med*; 342:1229-35.
15. Guruprasad K, Reddy BVP, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Prot Eng 4: 155-164
16. Harcourt B.H., L. Lowe, A. Tamin, X. Liu, et al. [2005] Genetic characterization of Nipah virus, Bangladesh, 2004. Emerg. Infect. Dis. 11:1594-1597.
17. Ikai AJ (1980) Thermo stability and aliphatic index of globular proteins. J Biochem 88: 1895-1898
18. Kyte J, Doolottle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105- 132.
19. Lamb RA, Parks GD.[2006] *Paramyxoviridae*: The Viruses and Their Replication. In: Knipe DM, Howley PM, editors. Fields Virology. Fifth ed. Philadelphia: Lippincott, Williams and Wilkins; pp. 1449–1496.
20. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8: 477-486
21. M. S. Chadha, J. A. Comer, L. Lowe, et al. [2006] Nipah virus-associated encephalitis outbreak, Siliguri, India. Emerg. Infect. Dis.12:235.
22. Mohd Nor MN, Gan CH, Ong BL. [2000] Nipah virus infection of pigs in peninsular Malaysia. *Rev Sci. Tech. Off Int Epiz*; 19:160-5.
23. Ramachandran GN, Ramakrishnan C, Sasisekhran V (1963) Stereochemistry of polypeptide chain confi guarations. J Mol Biol 7: 95-99.
24. Takimoto T, Portner [2004] A. Molecular mechanism of paramyxovirus budding. Virus Res. ;106:133–145.
25. Vriend G (1990) WHAT IF: A molecular modeling and drug design program. J Mol Graph 8: 52-56