JOURNAL OF RECENT ADVANCES IN APPLIED SCIENCES (JRAAS) 29:41-50, 2014
ISSN 0970 1990, ST CODE 08 ,www.jraas.org. E MAIL: mhcdr@rediff.com
DISPLAYED ON LINE MAY 2013

STRUCTURE PREDICTION AND FUNCTIONAL CHARACTERIZATION OF MATRIX PROTEINS OF HUMAN METAPNEUMOVIRUS (STRAIN CAN97-83) (HMPV)

Navneet Kumar Yadav · Pooja Shukla · *Shruti Pareek · **Ashok Kumar Srivastava · ***Falgun Wanganugi Bansode and #R. K. Singh

Divisions of Toxicology,**Computer and ***Endocrinology, CSIR-Central Drug Research Institute, Lucknow- 226001, Jaipur National University*, Jaipur India,
Corresponding Author: Mailing address: Division of Toxicology, CSIR – Central Drug Research Institute, Lucknow, India – 226001.Phone:919984941595. E – mail: rktox@yahoo.com

ABSTRACT
The pathogen for most of the acute respiratory illness [ARI] is a close relative of the avian pneumovirus, a member of the *Metapneumovirus* genus and was called Human Metapneumovirus [hMPV]. A bioinformatics and molecular modeling approach was adopted to explore properties and structure of matrix proteins of Human Metapneumovirus [strain CAN97-83] [hMPV]. The Matrix proteins selected for this study are MTRX_HMPVC, M21_HMPVC and M22_HMPVC. Physico-chemical characterization interprets properties such as isoelectric point [pI], extinction coefficient [EC], aliphatic index [AI], grand average hydropathy [GRAVY] and instability index [II] and provides data about these proteins and their properties. Prediction of motifs, patterns, disulfide bridges and secondary structure were performed for functional characterization. Three dimensional structures for these proteins are still not available in Protein Data Bank [PDB]. Therefore, homology model for selected Matrix protein is developed. The modelling of the three dimensional structure of this protein shows that model generated by Modeller was more acceptable in comparison to that by Geno3D and Swiss Model. The models were validated using protein structure checking tools PROCHECK, WHAT IF and Verify_3D. These structures will provide a good foundation for functional analysis of experimentally derived crystal structures. Finally selected Modelled structure was submitted to protein model database and can be downloaded using accession number PM0077331. The knowledge of the structure of the protein has a capability to enlighten the protein functions and its active sites, virtual screening, facilitating site-directed mutagenesis studies and drug design.

KEY WORDS: Homology Modeling, Human Metapneumovirus [hMPV], Physico-chemical characterization, Matrix proteins.

ABBREVIATIONS: ARI, acute respiratory illness; hMPV, Human Metapneumovirus; G, attachment; F, fusion; SH, small hydrophobic protein; N, nucleocapsid protein; P, phosphoprotein; L, large polymerase protein; M, matrix protein; pI, isoelectric point; EC, extinction coefficient; AI, aliphatic index; GRAVY, grand average hydropathy;II, instability index; PDB, Protein Data Bank ; ARI, Acute Respiratory Infections; RSV,Respiratory Syncytial Virus; APV, avian pneumovirus ; SOPMA, Self Optimized Prediction Method with Alignment; GROMACS, GROningen MAchine for Chemical Simulations;

INTRODUCTION
In the year 2000, millions of children below the age of 5 years died from Acute Lower Respiratory tract Infections [ALRI] across the world [1]. The leading causes of this worldwide mortality are the Acute Respiratory Infections [ARI] [2,3]. All the classes of micro-organisms including viruses, bacteria and protozoa are capable of causing respiratory tract infections; the most common are viruses and bacteria. Despite of extensive diagnostic testing, not a single pathogen was attributed for a substantial portion of respiratory infections.

The recently discovered pathogen for most of the Acute Respiratory Infections [ARI] is a Metapneumovirus. It was first reported in Netherlands in the year 2001 in 28 hospitalized children and infants having ARI with symptoms similar to that of Respiratory Syncytial Virus [RSV] [4]. The virus was found to be the close relative of the avian pneumovirus [5], a member of the *Metapneumovirus* genus and was called human metapneumovirus [hMPV] [6].

When the virus was genetically analyzed, it was found to be a RNA virus belonging to the family of *Paramyxoviridae* and subfamily of *Pneumovirinae.* The avian pneumovirus [APV] and hMPV are related to each other, and taxonomists separated these two into a separate genus, *Metapnemovirus* [7].

As revealed by electron microscopy, the virus is morphologically similar to paramyxovirus. A lipid envelope derived from the plasma membrane of the host cell surrounds the virion. The envelope has three virus glycoprotein inserted into it, the attachment [G], fusion [F], and small hydrophobic [SH] protein [8]. The RNA genome and the viral proteins are associated to form the helical nucleocapsid [represented on the right and in the centre of the virion below in [figure 1]. The proteins consist of the nucleocapsid protein [N], the phosphoprotein [P], and the large polymerase [L] protein. The M2-1 transcriptional enhancer protein is also thought to be associated with the nucleocapsid. The matrix [M] protein surrounds the nucleocapsid and forms a link between the nucleocapsid and the lipid membrane of the virus particle [9].

The genome size of hMPV virus is about 13 kb in length [9], it is a negative sense non-segmented RNA. The genome is predicted to encode 9 proteins in the order 3-N-PM- F-M2-SH-G-L-5, the M2 gene is predicted to encode 2 proteins, M2-1 and M2-2 [5] [figure 2]. The genome has a noncoding 3' leader which has the viral promoter, 5' trailer and intergenic regions [10] but they lack 2 nonstructural proteins which are present in RSV genomes. These proteins counteract with the host interferons so the absence of these genes in the metapneumoviruses may affect the relative pathogenicity of these viruses as compared to RSV strains [11]. The Sequence of APV and hMPV open reading frames are 56% to 88% identical [5]. The function of each of the gene products was predicted by comparison with other paramyxoviruses. The F [fusion], G [glycosylated] and SH [short hydrophobic] proteins are integral membrane proteins on the surface of infected cells and virion particles. The F protein appears to be a classic viral fusion protein, with a predicted nonfurin F1/F2 cleavage site near a hydrophobic fusion peptide and 2 heptad repeats in the extra cellular domain that facilitate membrane fusion. The predicted G protein of hMPV exhibits the basic features of a glycosylated type II mucin-like protein but interestingly lacks the cluster of conserved cysteines sometimes termed the "cysteine noose" that is found in the RSV and APV G proteins. G gene of hMPV tends to be highly variable like RSVG gene which may be due to host immune selection pressure [12]. The function of the SH proteins of both the viruses remains unknown. The N [nucleoprotein], P [phosphoprotein] and L [large, polymerase] proteins are replication proteins in the nucleocapsid, the M2-1 and M2-2 proteins are regulatory proteins and the M [matrix] protein may coordinate viral assembly of viral nucleocapsids with envelope proteins [12].

MATRIX PROTEIN

MTRX_HMPVC has a crucial role in virus assembly and budding. The matrix interacts with the RNP complex and this association serves two functions: facilitate virion assembly and inhibit the viral transcriptase activity. Early in infection, M is localized to the nucleus and may inhibit host cell transcription. Later on, M can associate with lipid rafts supposedly by interacting with the cytoskeleton and with the cytoplasmic tail of glycoprotein G. The binding of M to host membrane is stabilized by the surface expression of the viral glycoproteins. These interactions may allow virus formation by mediating association of the nucleocapsid with the nascent envelope [13].

M21_HMPVC acts as a transcriptional elongation factor to prevent premature termination during transcription thus allowing complete synthesis of viral mRNAs. It functions also as a processivity and antitermination factor to permit transit of the polymerase through intergenic regions to access promoter distal genes. It also plays a role in the association of the matrix protein with the nucleocapsid, which initiates assembly and budding [13].

M22_HMPVC mediates the regulatory switch from transcription to RNA replication. It acts late in infection by inhibiting viral transcription and up regulating RNA replication [13].

Experimental determination of protein structure through X-ray crystallography or nuclear magnetic resonance spectroscopy is time consuming and is a very costly affair. Protein Data Bank is a repository for three-dimensional structural data of large biological molecules submitted by biologists and biochemists from around the world. Still, majority of protein sequences have no structural information as the number of unique structural folds that proteins adopt is limited and number of experimentally determined new structures is increasing exponentially. Therefore, it is an obvious demand to bridge this 'structure knowledge gap' and computational methods for protein structure prediction. Various approaches have been followed in this context. The three-dimensional

structures of proteins in a family are obvious to be conserved more than their sequences. Hence, if one detects the similarity between two proteins at the sequence level, structural similarity can usually be assumed further. Moreover, it may happen that proteins that have no detectable sequence similarity may have similar structures. Many structure-function relationships can be deduced from a reasonable model, which may further be used for successful drug design. Considering all these aspects, homology modeling of the Matrix Proteins of Human Metapneumovirus [hMPV] [Strain CAN97-83] under study is performed.

MATERIALS AND METHODS

Sequences of Matrix Proteins of Human Metapneumovirus [hMPV] [Strain CAN97-83] are retrieved from the SWISSPROT, a public domain protein database [14]. Table-1 shows the protein sequences considered in this study. The sequences are retrieved in FASTA format and used for further analysis.

PHYSICO-CHEMICAL CHARACTERIZATION

For physico-chemical characterization, theoretical isoelectric point, molecular weight, total number of positive and negative residues, extinction coefficient [15], instability index [16], aliphatic index [17] and grand average hydropathy [18] are computed using the Expasy's ProtParam server [19] [*http://us.expasy.org/tools/protparam.html*]. The results are shown in Table 2.

FUNCTIONAL CHARACTERIZATION

The SOSUI server performed the identification of membrane and soluble proteins from amino acid sequence. Disulphide bonds are important in determining the functional linkages. Table 3 shows the prediction of "SS" bonds using the primary structure [protein sequence data] by the tool CYS_REC [http://linux1.softberry.com/berry.phtml?topic=cys_rec&group=programs&subgroup=propt].
CYS_REC identifies the position of cysteins, total number of cysteins present and pattern, if present, of pairs in the protein sequence. Prosite is a database of protein families and domains [20]. Table 4 represents the output of Prosite that is recorded in terms of the length of amino residues of protein with specific profiles and patterns.

SECONDARY STRUCTURE PREDICTION:

SOPMA [21] is employed for calculating the secondary structural features of the antioxidant protein sequences considered for this study. The results are presented in Table 5.

SEQUENCE ALIGNMENT , MODEL BUILDING AND EVALUATION:

```
sp|Q6WB99|MTRX_HMPVC                                                          -
MESYLVDTYQGIPYTAAVQVDLVEKDLLPASLTIWFPLFQANTPPAVLL 49
gi|223673799|pdb|2VQP|A EMETYVNKLHEGSTYTAAVQYNVLEKDDDPASLTIWVPMFQSSMPADLLI 50
                        **:*:  . .:* .****** :::***  *******.*:**:. *. .*:
sp|Q6WB99|MTRX_HMPVC
DQLKTLTITTLYAASQSGPILKVNASAQGAAMSVLPKKFEVNATVALDEY 99
gi|223673799|pdb|2VQP|A KELANVNILVKQISTPKGPSLRVMINSRSAVLAQMPSKFTICANVSLDDR 100
                        .:* .:.* .   :: .** *:*  .::.*.:: :*.** : * *:**:
sp|Q6WB99|MTRX_HMPVC
SKLEFDKLTVCEVKTVYLTTMKPYGMVSKFVSSAKPVGKKTHDLIALCDF 149
gi|223673799|pdb|2VQP|A  SKLAYDVTTPCEIKACSLTCLKSKNMLTTVKDLTMKTLNPTHDIIALCEF 150
                        *** :*  * **:*:  ** :*. . *::.. . :  . : ***:*****:*
sp|Q6WB99|MTRX_HMPVC  MDLEKNTPVTIPAFIKSVSIKESESATVEAAISSEADQALTQAKIAPYAG 199
gi|223673799|pdb|2VQP|A     ENIVTSKKVIIPTYLRSISVRNKDLNTLENITTTEFKNAITNAKIIPYSG 200
                        :: ...* **::::*:*:::. :  *:*   ::* .:*:*:*** **:*
sp|Q6WB99|MTRX_HMPVC
LIMIMTMNNPKGIFKKLGAGTQVIVELGAYVQAESISKICKTWSHQGTRY 249
gi|223673799|pdb|2VQP|A LLLVITVTDNKGAFKYIKPQSQFIVDLGAYLEKESIYYVTTNWKHTATRF 250
                        *::::*:.:: ** ** :  . :*.**:****:: ***  : ..*.* .**:

sp|Q6WB99|MTRX_HMPVC        VLKSR-- 254
gi|223673799|pdb|2VQP|A     AIKPRED 257
                        .:*.*
```

Sequence alignment of target sequence is performed in BLASTp, Structure Of The Matrix Protein From Human Respiratory Syncytial Virus [PDB id 2VQP_A 38% identity,63% similarity and 100% Query coverage] is selected as template sequence on the basis of sequence identity and query coverage with target sequence.

The modelling of the three dimensional structure of the protein is performed by three homology modelling programs, Geno 3D [22], Swiss Model [23] and Modeller [24]. The overall stereochemical property of the protein is assessed by Ramchandran plot analysis [25]. The validation for structure models obtained from the three software tools is performed by using PROCHECK [26]. The models are further checked with WHAT IF [27] and Verify_3D. The results of PROCHECK, WHAT IF and Verify_3D analysis are shown in Table 6, Table 7 and Table 8 respectively. Structural analysis is performed and figures representations are generated with Swiss PDB Viewer [29].

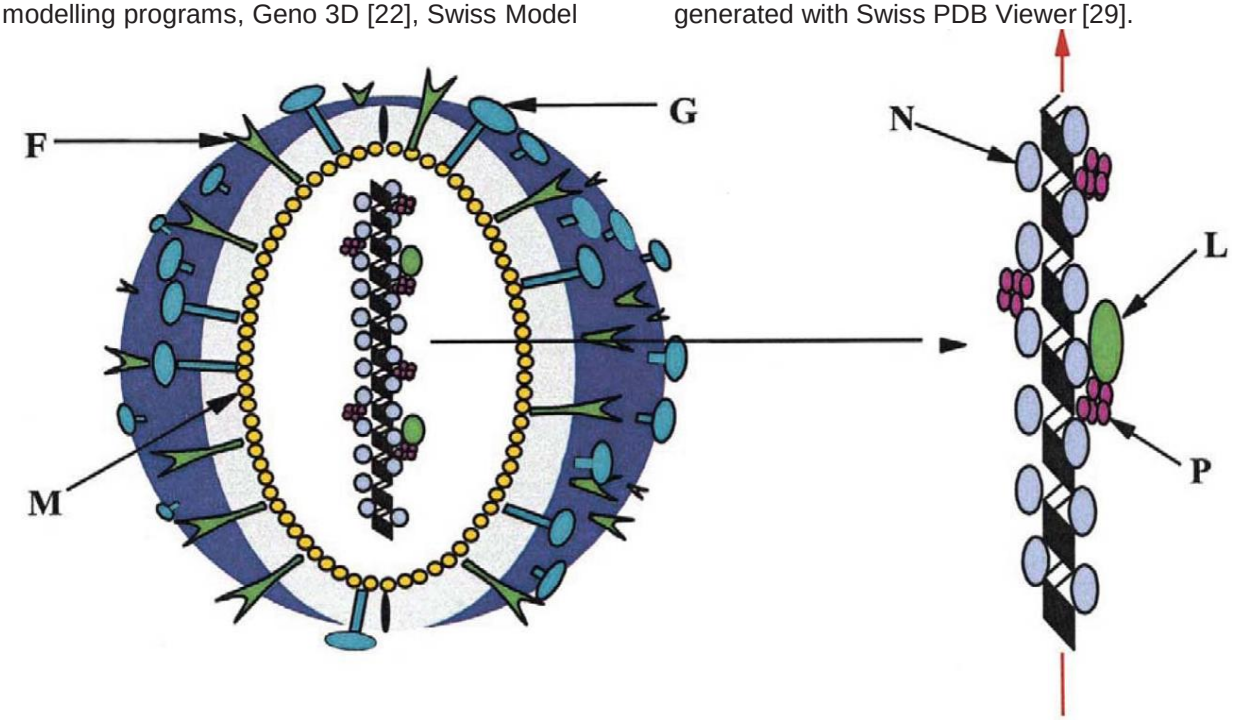

Figure 1: Schematic diagram of the pneumovirus particle (Source: Easton, et al, 2004).

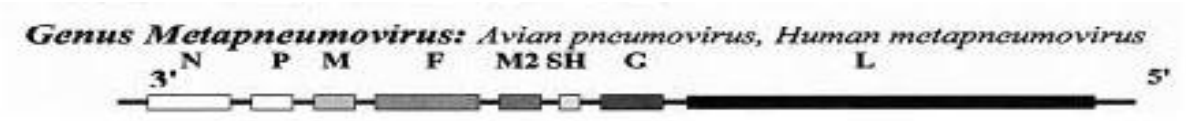

Figure 2: Negative-stranded RNA linear genome, about 13 kb in size, Encodes for eight proteins (Source: Easton et al, 2004)

Table-1: Protein sequences considered for the study

Matrix Proteins	Accession No.	Length
MTRX_HMPVC	**Q6WB99**	254
M21_HMPVC	**Q6WB97**	**187**
M22_HMPVC	**Q6WB96**	71

Table-2: Parameters computed using Expasy's ProtParam tool

Matrix Proteins	Accession No.	Length	M.wt	pI	-R	+R	EC	II	AI	GRAVY
MTRX_HMPVC	**Q6WB99**	254	27612.2	8.27	23	25	26025	16.56	98.70	0.144
M21_HMPVC	**Q6WB97**	**187**	21234.0	9.14	21	27	22920	46.27	90.27	-0.582
M22_HMPVC	**Q6WB96**	71	8250.9	6.04	9	8	6990	25.35	130.28	0.43

Table-3: Disulphide (SS) bond pattern of pairs predicted, by CYS_REC

Matrix Proteins	Accession No.	Length	CYS_REC
MTRX_HMPVC	**Q6WB99**	254	_
M21_HMPVC	**Q6WB97**	**187**	Cys7-Cys21
M22_HMPVC	**Q6WB96**	71	_

Table-4: Functional characterization of proteins of spinach at Prosite

Matrix Proteins	Accession No.	Profile	Position in the protein	Description
MTRX_HMPVC	**Q6WB99**	**-**	**-**	**-**
M21_HMPVC	**Q6WB97**	ZF_C3H1	**1-28**	Contains a zinc-finger domain on its N-terminus essential for its antitermination function
M22_HMPVC	**Q6WB96**	**-**	**-**	**-**

Table-5: Calculated secondary structure elements by SOPMA

Matrix Proteins	MTRX_HMPVC	M21_HMPVC	M22_HMPVC
Secondary structure	**Q6WB99**	**Q6WB97**	**Q6WB96**
Alpha helix	33.07%	44.92%	59.15%
3_{10} helix	0.00%	0.00%	0.00%
Pi helix	0.00%	0.00%	0.00%
Beta bridge	0.00%	0.00%	0.00%
Extended strand	30.31%	16.58%	23.94%
Beta turn	7.48%	3.74%	5.63%
Bend region	0.00%	0.00%	0.00%
Random coil	29.13%	34.76%	11.27%
Ambigous states	0.00%	0.00%	0.00%
Other states	0.00%	0.00%	0.00%

Table-6: Ramachandran plot calculation and Comparative analysis of the models from Geno3D, Swiss-model and Modeller computed with the PROCHECK program

Matrix Protein	Server		Geno3D	Swiss model	Modeller
MTRX_HMPVC	PDB templates Used	PDB code	2VQP_A	2VQP_A	2VQP_A
		Accession number	P03419	P03419	P03419
	Residues in the most Favored Region		79.4%	86.8%	93.4%
	Residues in additionally allowed region		18.9%	12.3%	6.6%
	Residues in generously allowed region		1.3%	0.9%	0.0%
	Residues in disallowed region		0.4%	0.0%	0.0%

Table-7: RMS Z-score for bond angles of modelled protein structure using WHAT IF.

Matrix Protein	Server	RMS Z-score for bond angles
MTRX_HMPVC	**Geno3D**	**0.404**
	Swiss model	**1.003**
	Modeller	**1.213**

Table-8: Verify_3D: Average 3D-1D score > 0.2

Matrix Protein	**Server**	Verify_3D
		the residues have an averaged 3D-1D score > 0.2
MTRX_HMPVC	**Geno3D**	89.80%
	Swiss model	94.51%
	Modeller	93.73%

RESULT AND DISCUSSION

Table-1 shows the Matrix Proteins of Human Metapneumovirus considered in this study. These Matrix Proteins sequences are retrieved from the SWISSPROT, a public domain protein database. These protein sequences are retrieved in FASTA format and used for further analysis. Parameters computed using Expasy's ProtParam tool is represented in Table-2.

The calculated isoelectric point [pI] is significant because at pI, solubility is least and mobility in an electro focusing system is zero. Isoelectric point is the pH at which the surface of protein is covered with charge but net charge of protein is zero. At their respective pI proteins are stable and compact. The computed pI value of M22_HMPVC is less than 7 [pI<7] indicates that this Matrix protein is acidic. The pI of MTRX_HMPVC and M21_HMPVC are greater than 7 [pI > 7] reveals that these proteins are basic in character. Practically, pI is useful for developing buffer system for purification by isoelectric focusing method.

Although Expasy's ProtParam computes the extinction coefficient for 276, 278, 279, 280 and 282 nm wavelengths, 280 nm is favored because proteins absorb light strongly at this locus while other substances do not. Extinction coefficient of matrix proteins at 280 nm is ranging from 6990 to 26025 M^{-1} cm^{-1} with respect to the concentration of Cys, Trp and Tyr. The high extinction coefficient of MTRX_HMPVC and M21_HMPVC indicates presence of high concentration of Cys, Trp and Tyr. The computed extinction coefficients help in the quantitative study of protein–protein and protein–ligand interactions in solution.

The instability index provides an estimate of the stability of protein in a test tube. There are certain dipeptides, the occurrence of which is significantly different in the unstable proteins compared with those in the stable ones. This method assigns a weight value of instability. Using these weight values it is possible to compute an instability index. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable [16]. The instability index value for the matrix proteins were found to be ranging from 16.56 to 46.27. The result shows MTRX_HMPVC and M22_HMPVC are stable proteins [29]

The aliphatic index [AI] which is defined as the relative volume of a protein occupied by aliphatic side chains [A, V, I and L] is regarded as a positive factor for the increase of thermal stability of globular proteins. Aliphatic index for the Matrix protein sequences ranged from 90.27 – 130.28. The very high aliphatic index of all Matrix protein sequences indicates that these Matrix proteins are stable for a wide temperature range.

The Grand Average hydropathy [GRAVY] value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence. GRAVY of Matrix proteins are ranging from -0.582 to 0.432. This low range of value indicates the possibility of better interaction with water. Functional analysis of these proteins includes prediction of transmembrane region, disulfide bond and identification of important motifs.

SOSUI distinguishes between membrane and soluble proteins from amino acid sequences, and predicts the transmembrane helices for the former. The server SOSUI classifies all Matrix proteins as soluble proteins. As disulphide bridges play an important role in determining the thermostability of these proteins. CYS_REC is used to determine the Cysteine residues and disulphide bonds. Possible pairing and pattern

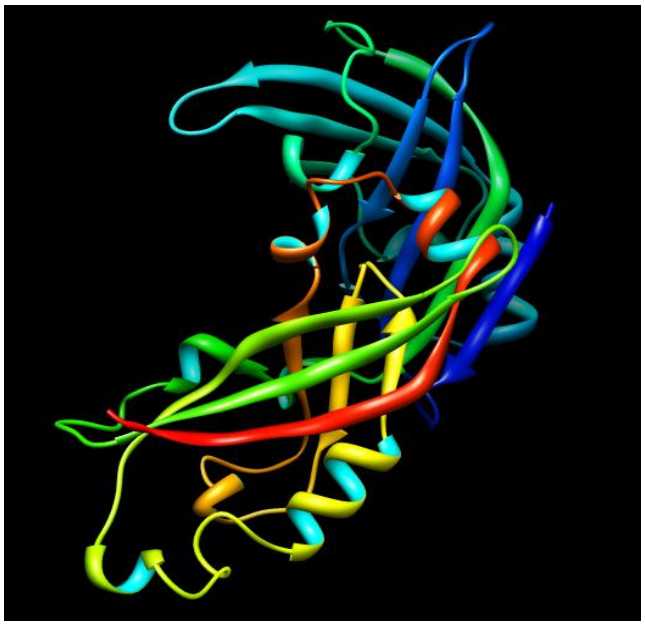

FIGURE 3: Modeled Structure of Matrix protein: MTRX_HMPVC

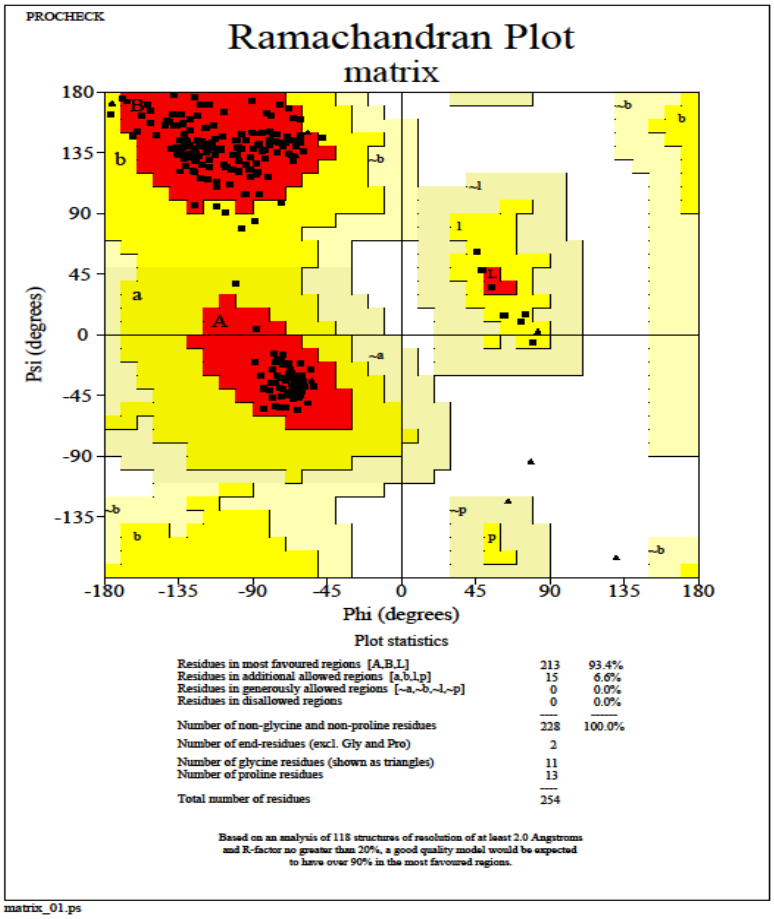

FIGURE 4: Ramachandran's Map of Matrix protein: MTRX_HMPVC with probability are presented in Table 3. Result shows that only M21_HMPVC protein contains disulphide linkages.

The functions of Matrix proteins are analyzed by submitting the amino acid sequence to Prosite server. Sequence of a particular cluster of residue types, which is variously known as a pattern, motif, signature or fingerprint. These motifs, typically around 10 to 20 amino acids in length, arise because specific residues and regions thought or proved to be important to the biological function of a group of proteins are conserved in both structure and sequence during evolutionary process [30]. Prosite analysis suggested the functionality of these proteins with profiles and patterns identified for characteristic functionality were represented in Table 4. The secondary structure of Matrix proteins are predicted by SOPMA [Self Optimized Prediction Method with Alignment] which correctly predicts 69.5% of amino acids for a state description of the secondary structure prediction [21]. The secondary structure indicates whether a given amino acid lies in a helix, strand or coil. Secondary structure features as predicted using SOPMA are represented in Table 5. The results revealed that alpha helix dominated among secondary structure elements for all sequences. The secondary structure are predicted by using default parameters [Window width: 17, similarity threshold: 8 and number of states: 4].

Three dimensional structures are predicted for proteins where such data is unavailable. There is lack of experimental structures for these proteins considered. Out of three matrix protein sequences three dimensional structure was modulated only for MTRX_HMPVC. The modeling of the three dimensional structure of the protein was performed by three homology modeling programs, Geno 3D, Swiss Model and Modeller. For constructed three dimensional models, energy minimized in GROMACS force field using steepest descent minimization Algorithms. The phi and psi distribution of the Ramachandran Map generated by non glycine, non proline residues were summarized in Table 7. A comparison of the results obtained from the Geno 3D, Swiss Model and Modeller, three different software tools in Table 7 shows that the models generated by Modeller are more acceptable in comparison to that by Geno3D and Swiss Model. The final modeled structures are visualized by Swiss PDB Viewer that is shown in figure 3.

The stereo chemical quality of the predicted models and accuracy of the protein model is evaluated after the refinement process using Ramachandran Map calculations computed with the PROCHECK program. The assessment of the predicted models generated by modeller is shown in Figure 4. The main chain parameters plotted on Ramachandran plot are quality, peptide bond planarity, bad non-bonded interactions, main chain hydrogen bond energy, C-alpha chirality and over-all G factor. In the Ramachandran plot analysis, the residues were classified according to its regions in the quadrangle. The red regions in the graph indicate the most allowed regions whereas the yellow regions represent allowed regions. Glycine is represented by triangles and other residues are represented by squares. The result revealed that the modeled structure for MTRX_HMPVC has 93.4% residue in allowed region. The distribution of the main chain bond lengths and bond angles are found to be within the limits for these proteins. Such figures obtained by Ramachandran plot represent a good quality predicted models. The modeled structures of matrix protein is also validated by other structure verification servers WHAT IF and Verify_3D. Standard bond angle of the model was determined using WHAT IF. The results are shown in Table 8. The analysis revealed RMS Z-scores are almost equal to 1 suggesting high model quality. The averaged 3D-1D score > 0.2 of the residues determined using Verify_3D were shown in Table 9. The predicted structures conforms the stereochemistry, indicating reasonably good quality. Modelled structure is submitted to protein model database which is a repository for three dimensional protein models obtained by structure prediction methods. Submitted, Modelled Matrix protein MTRX_HMPVC can be downloaded from PMDB using accession number PM0077331.

CONCLUSION

Homology modelling of matrix protein MTRX_HMPVC is performed. MTRX_HMPVC has a crucial role in virus assembly and budding, Homology derived model have a wide range of applications such as virtual screening, drug design, site-directed mutagenesis experiments or in rationalizing the effects of sequence variation [31]. Understanding the structure of proteins reveals the protein function and active sites .These structures will serve as cornerstone for functional analysis of experimentally derived crystal structures. [32,33].

REFERENCES

1. Broor S, Bharaj P, Chahar HS (2008) Human metapneumovirus: a new respiratory . pathogen. J. Biosci. 33:483–493.
2. Denny FW, Loda FA (1986) Acute respiratory infections are the leading cause of death in children in developing countries. Am. J. Trop. Med. Hyg. 35:1– 352.
3. Hijazi Z, Pacsa A, Eisa S, Shazli Ael, Abd el-Salam RA (1996) Laboratory diagnosis of acute lower respiratory tract viral infections in children. J. Trop. Pediatr. 42:276–280.
4. Van den Hoogen BG, De Jong JC, Groen J, Kuiken T, De Groot R, Fouchier RA, Osterhaus AD (2001) A newly discovered human pneumovirus isolated from young children with respiratory tract diseases. Nat. Med. 7:719–724.
5. Van den Hoogen BG,. Bestebroer TM, Osterhaus AD, Fouchier RA (2002) Analysis of the genomic sequence of a human Metapneumovirus. Virology 295:119–132.
6. Van den Hoogen BG, Herfst S, Sprong L, Cane PA, Forleo-Neto E, De Swart RL, Osterhaus AD, Fouchier RA (2004) Antigenic and genetic variability of human metapneumoviruses, Emerg. Infect. Dis. 10:658–666.
7. Lamb RA, Collins PL, Kolakofsky D, Melero JA, Nagai Y, Oldstone MBA (2000) Family paramyxoviridae. In: Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Maniloff J (eds.) Virus taxonomy, Academic Press, San Diego, pp. 549–561.
8. Collins PL, Mottet G (1993) Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory Syncytial virus. J. Gen. Virol. 74:1445–1450
9. Easton AJ, Domachowske JB, Rosenberg HF (2004) Animal pneumovirus: molecular genetics and pathogenesis. Clin. Microbiol. Rev. 17:390–412.
10. Herfst S, De Graaf M, Schickli JH, Tang RS, Kaur J, Yang CF, Spaete RR, Haller AA, Van den Hoogen BG , Osterhaus AD, Fouchier RA (2004) Recovery of human metapneumovirus genetic lineages A and B from cloned cDNA. J. Virol. 78: 8264–8270.
11. Collins PL, Chanock RM, Murphy BR (2001) Respiratory Syncytial virus. In : Knipe DM,. Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (eds.) Fields virology, 4th edition, Lippincott-Raven Publishers, Philadelphia,Vol.1, pp. 1443–1485.
12. Crowe EJ (2004) Human Metapneumovirus as a Major Cause of Human Respiratory Tract Disease. Pediatr. Infect. Dis. J. 23:S215–S221.
13. Pham QN, Biacchesi S, Skiadopoulos MH, Murphy BR, Collins PL, Buchholz UL (2005) Chimeric recombinant human metapneumoviruses with the nucleoprotein or phosphoprotein open reading frame replaced by that of avian metapneumovirus exhibit improved growth in vitro and attenuation in vivo. J. Virol. 79:(24)15114-15122.
14. MISSING
15. Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and it supplement TrEMBL in 2000. Nucleic Acids Res. 28:45-48.
16. Gill SC, Von Hippel PH (1989) Extinction coefficient. Anal Biochem. 182:319-328.
17. 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.
18. Ikai AJ (1980) Thermo stability and aliphatic index of globular proteins. J Biochem. 88:1895-1898.
19. Kyte J,. Doolottle RF (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.
20. Gasteiger E, (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: M.J. Walker (eds.) The Proteomics Protocols Handbook, Humana Press, New York, pp. 571-607.
21. Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJ, Hofmann K, Bairoch A (2002) The PROSITE database, its status in 2002. Nucleic Acids Res. 30:235-238.
22. Geourjon C, Deléage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Appl. Biosci. 11:681-684.
23. Combet C, Jambon M, Deleage G, Geourjon C (2002) Geno3D: Automatic comparative molecular modelling of protein. Bioinformatics 18:213-214.
24. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web based environment for protein structure homology modeling. Bioinformatics 22:195-201.
25. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE,. Berendsen HJ (2005) GROMACS: Fast, Flexible and Free. J. Comp. Chem. 26:1701-1718.
26. Ramachandran GN, Ramakrishnan C, Sasisekhran V (1963) Stereochemistry of polypeptide chain configuarations. J. Mol. Biol. 7:95-99.
27. 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.

28. Vriend G. (1990) WHAT IF: A molecular modeling and drug design program. J. Mol. Graph 8:52-56.
29. Guex N, Manuel CP (1997) Data Modeling, Analysis and Classification SWISSMODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis 18:2714-2723.
30. Sahay A, Shakya M (2010) In silico Analysis and Homology Modelling of Antioxidant Proteins of Spinach. J. Proteomics Bioinform. 3:148-154.
31. Sigrist CJ, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, Bairoch A, Bucher P (2002) PROSITE: A documented database using patterns and profi les as motif descriptors. Briefings in Bioinformatics 3:265-274.
32. Roy S, Maheshwari N, Chauhan R, Sen NK, Sharma A, (2011) Structure prediction and functional characterization of secondary metabolite proteins of *Ocimum* Bioinformatics 6:(8)315-319.
33. Yadav NK, Sarika, Iquebal MA, Akram M, (2011) In silico analysis and homology modeling of coat protein of Mungbean Yellow Mosaic Indian Virus Journal of Food Legumes 24:(2)138-141.