*International Journal of Food, Agriculture and Veterinary Sciences ISSN: 2277-209X (Online)*
*An Online International Journal Available at http://www.cibtech.org/jfav.htm*
*2012 Vol. 2 (3) September-December, pp. 26-36/Ashokan et al.*
***Research Article***

# SEQUENCE ANALYSIS OF PROTEIN IN PESTE DES PETITS RUMINANTS: AN *IN SILICO* APPROACH

**Munduganore D.S.[1], Mundaganore Y.D.[2] and *Ashokan K.V.[3]**
*[1]Department of Zoology, Willingdon College, Sangli,MS-416416*
*[2]Department of Zoology, MIraj College , Miraj, M.S-416410*
*[3]Department of Zoology, P.V.P.College, Kavathe Mahankal, Sangli,M.S-416405*
**Author for Correspondence*

**ABSTRACT**
This investigation aim to determine the molecular identity of various proteins of PPR so as to pave the way to find out better vaccine and therapeutic method to treat and control the devastating disease of small ruminants. The protein sequence analysis reveals that most of the proteins are rich in leucine with some exemptions that have glycine predominantly. The absence of tryptophan in nucleocapsid protein and fusion protein indicate the requirement of sophisticated method for its extraction and purification. The proteins of PPR appear better stability, an indication of resistance to mutation. A further study on mutation considering this view is needed. The stability of the protein is also obvious from the secondary structure study, 3D model generated and its validation data including Ramachandran plot. The transmembrane domain is observed only in two proteins (Fusion protein and haemagglutinin protein), these protein may be the candidate helping for entering the virus in to the hoist cell besides its normal function.

***Key Words:*** *Peste Des Petits Ruminants, ProtParam, SOPMA, SVMProt, SOSUI, VADAR and Ramchandran Map*

## INTRODUCTION

Peste des petits ruminants (PPR) is a highly contagious viral disease of small ruminants caused by *Morbillivirus* belongs to the family Paramyxoviridae. The virus is closely related to rindrepest virus of bovines and buffaloes, distemper virus of dogs and other wild carnivores, human measles virus and Morbilliviruses of marine mammals (Barrett *et al.*, 1993., Jones *et al.*, 1993., Scott, 1981, Yayehard, 1997). Four genetic lineages (lineage 1-40) of this virus have been identified. The severity of the disease varies with species, as well as the animals' immunity to PPRV and its breed. Morbidity in the range of 10 to 80% and mortality proportions from 0 to 90% has been reported. The wide range of reported values is likely to be influenced by differences between species (sheep or goats), production systems and levels of natural or acquired immunity (Akakpo *et al.*, 1996, Diallo, 2003, Nanda *et al.*, 1996, Pegram and Tereke, 1981., and Wakwaya, 1997). PPR was first described in West Africa in 1942 (Gargadennec and Lalanne, 1942). Nowadays the disease is recognized as responsible for mortality and morbidity across most of the sub-Saharan African countries situated north of the equator, in the Arabian Peninsula, in India and in numerous other countries in Asia (Diallo, 2003., Abraham, 2005, Shaila, 1996). Now it is found that several sheep in south India suffering from rinderpest are actually infected with peste des petits ruminants (PPR thought to be restricted to West Africa. Rinderpest usually afflicts Indian cattle, wild Indian bison and smaller ruminants such as sheep and goats. PPR is found more often among goats and sheep. It is observed that PPR-infected animals in India suffer from ulcers in the respiratory tract, high fever, diarrhoea and eventually die. In India, so far, PPR has affected only sheep. The PPR outbreak in India was first reported in 1989 by M S Shaila of the department of microbiology and cell biology at Bangalore's Indian Institute of Science, when about 80 sheep in a flock of 800 at Arasur village in Villupurum district of Tamil Nadu were affected. Last year, PPR outbreaks were reported and confirmed from Karnataka, Tamil Nadu and Andhra Pradesh, says the scientist. Shaila and her collaborators from the Madras Veterinary College used special diagnostic probes to identify the PPR virus, which are usually

difficult to tell apart from the rinderpest virus. Experiments have shown that vaccinating goats and sheep against rinderpest can protect them against PPR, too. However, according to Shaila, it is not known how effective the immunity provided by the rinderpest vaccination will be in the long run. She explains, "This is a new disease in India and efforts must be made to ensure that it does not spread to other parts of the country."

The virus is excreted in oculonasal discharges, saliva and feces at the onset of the clinical signs (Scot, 1981). The usual incubation period is from two to six days before fever and mucosal erosions occur. Diarrhea develops two to three days later and death is usually preceded by pneumonia (Hamdy *et al.*, 1976). There is no specific treatment for PPR, however hyper immune PPR serum produced in goats reverses the disease process if administered at the onset of fever (Themelandu, 1985). It can be prevented by attenuated RPV vaccination (Taylor, 1979).

The purpose of this investigation was to determine the molecular identity of various proteins of PPR so as to pave the way to find out better vaccine and therapeutic method to treat and control the devastating disease of small ruminants.

## MATERIALS AND METHODS

The protein sequence of various proteins of PPR identified so far is retrieved from NCBI databank (Table 1). The sequences were used for further characterization of protein and generation of 3D structure and its validation.

**Table1: Various protein sequences mined from NCBI data bank**

| Protein | Accession Number | Number of amino acid |
|---|---|---|
| Fusion Protein | ACF24469.1 | 149 |
| F-Protein | AAV76020.1 | 125 |
| NucleoProtein | ABY61077.1 | 116 |
| L-protein | ABY71272.1 | 2183 |
| H-Protein | ABY71271.1 | 609 |
| Nucleo Capsid Protein | ABA03161.1 | 94 |
| C-protein | ACQ44674.1 | 177 |
| V-Protein | ACQ44673.1 | 298 |
| Large Protein | ACQ44672.1 | 2183 |
| Haeagglutinin | ACQ44671.1 | 609 |
| Matrix protein | ACQ44669.1 | 335 |
| Phospho Protein | ACQ44668.1 | 509 |
| M-Protein | ABX75302.1 | 335 |
| P-Protein | ABX75300.1 | 509 |
| N-Protein | AAS68026.1 | 525 |
| RNA dependent RNA polymerase | CAJ01701.1 | 2183 |
| Non Structural Protein | CAJ01697.1 | 177 |
| Hypothetic Protein I | YP_133823.2 | 398 |
| Hypothetic Protein II | YP_133824.1 | 177 |
| Polymerase | CAD88265.2 | 2123 |

The amino acid sequence was first subjected for ProtParam Tool for computation of various physical and chemical parameters. The computational parameters include molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half life, instability index, aliphatic index and grand average of hydropathy (GRAVY) (Gasteiger *et al.*, 2005). The sequences were further subjected for secondary structure prediction by ExPASy's SOPMA tool. SOPMA **i**s an improved SOPM method. It predicts 69.5% of amino acids for a 3 state description of the secondary structure (a helix, b sheets and coil). It predicts the secondary structure by consensus prediction from multiple alignments.

To find out the functional role of the protein is essentially necessary to assign the protein's function in the drug resistance and infectivity. SVMProt tool was used to predict the functional signature of the selected protein sequences. SVMProt classify the protein into its functional family from its primary sequence (Cai

*et al.,* 2003). SVMProt classification system is trained from representative proteins of a number of functional families and seed proteins of Pfam curated protein families. The computed accuracy for protein family classification is found to be in the range of 69.1-99.6%. SVMProt shows a certain degree of capability for the classification of distantly related proteins and homologous proteins of different function and thus may be used as a protein function prediction tool that complements sequence alignment methods (Cai *et al.,* 2003).

The entry of bacteria in to a cell is a pre-requisite to attain the full infectivity of the bacteria. Transmembrane protein plays an important role in this functional role. The sequences were thus subjected to transmembrane domain identification by SOSUI server (Mitaku *et al.,* 2003). The system SOSUI was developed for the discrimination of membrane proteins and soluble ones together with the prediction of transmembrane helices, in which the accuracy of the classification of proteins was 99% and the corresponding value for the transmembrane helix prediction was 97% (Mitaku *et al.,* 2003).

To assign the function more precisely we subjected the sequences for 3D model generation by SWISS model server (Arnold *et al.,* 2006., Schwede *et al.,* 2003., Guex and Peitsch, 1997). The SWISS-MODEL Workspace is a web-based integrated service dedicated to protein structure homology modeling. It assists and guides the user in building protein homology models at different levels of complexity.

The 3D structure was then validated by using VADAR server and ramachandran plot. VADAR (Volume, Area, Dihedral Angle Reporter) is a compilation of more than 15 different algorithms and programs for analyzing and assessing peptide and protein structures from their PDB coordinate data. The results have been validated through extensive comparison to published data and careful visual inspection. The VADAR web server supports the submission of either PDB formatted files or PDB accession numbers. VADAR produces extensive tables and high quality graphs for quantitatively and qualitatively assessing protein structures determined by X-ray crystallography, NMR spectroscopy, 3D-threading or homology modeling (Leigh, *et al.,* 2003).

## RESULT AND DISCUSSION

The physical and chemical parameters show that some of the proteins share common parameters. This helps us to classify the proteins into smaller groups. The fusion protein and F-protein are one and same (Fusion protein here after) with some diversion may be due to the variation in the length of amino acid s taken for the study. Thus for nucleo-protein and N-protein (Nucleo protein), large protein and L-protein (large protein here after), RNA dependent RNA polymerase and Polymerase protein (Polymerase here after), H-protein and haemagglutinin protein (Haemagglutinin here after), matrix protein and M-protein (Matrix protein here after) and P-protein and phosopho protein (Phospho protein here after). The fusion glycoprotein (F protein) of paramyxoviruses plays a vital role in virus-induced cytopathology. The purified F protein, when incubated with chicken erythrocytes, caused lysis suggesting that PPRV F protein is a hemolysin (Devireddy *et al.,* 1999). We observed that the fusion protein is rich in leucine amino acid (10-12%). As Leucine is an aliphatic, hydrophobic, amino acid, it prefers substitution with other amino acids of the same type. Being hydrophobic, Leucine prefers to be buried in protein hydrophobic cores. It also shows a preference for being within alpha helices more so than in beta strands. The Leucine side chain is very non-reactive, and is thus rarely directly involved in protein function, though it can play a role in substrate recognition. In particular, hydrophobic amino acids can be involved in binding/recognition of hydrophobic ligands such as lipids (Betts, R.B. Russell, 2003). The same trend is observed in all the proteins selected except nucleo protein, nucleo capsid protein, P-protein, V-protein and hypothetical protein-I, here glycine is predominant (10-11%). It is unique among the proteinogenic amino acids in that it is not chiral. It is involved in the formation of many proteins and may takes part in the regeneration of the viral structural proteins. The various physical parameters of selected proteins are important for chemist, drug developers and molecular biologists significantly. The pI of the protein selected showed a range of (5-10) %.Isoelectric point is the pH at which a protein carries no net charge. The isoelectric point is of significance in protein purification because it is the pH at which solubility is

often minimal and at which mobility in an electro focusing system is zero (and therefore the point at which the protein will accumulate. This measure indicates how much light is absorbed by a protein at a particular wavelength. The extinction coefficient of nucleo protein is zero; this fact is reflected from the amino acid percentage study having no tryptophan in nucleo protein. The extinction coefficient of a protein at 280 nm depends almost exclusively on the number of aromatic residues, particularly tryptophan. Thus the purification of nucleoprotein needs an alternative method from other proteins in PPR. Knowing the extinction coefficient, the absorbance (optical density) can be calculated and subsequently the concentration.

**Table2: Amino AcidComposition**

| F- Protein % | L protein % | H Protein % | Nucleo capsid % | C Protein % | V Protein % | Large Protein % | Haemo agglutinin % | Matrix protein % | Phospho Protein % | M Protein % | P Protein % | N Protein % | RNA -RNA Polymerase % | Non structural protein % | Hypothetic % | Hypothetic protein % | Polymerase % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 5 | 3.8 | 6 | 6 | 5 | 5 | 4 | 4.5 | 6 | 4.5 | 5.5 | 9 | 5 | 5 | 6.4 | 5.6 | 5.4 |
| 6 | 6 | 7.1 | 12 | 8 | 6 | 6 | 7 | 7.2 | 4 | 7.5 | 3.9 | 8 | 6 | 7 | 4.7 | 8.5 | 5.7 |
| 5 | 4 | 4.3 | 2 | 2 | 5 | 4 | 4 | 5.4 | 4 | 5.4 | 3.3 | 3 | 4 | 2 | 4.4 | 1.7 | 4.2 |
| 2 | 5 | 5.9 | 4 | 3 | 9 | 5 | 6 | 5.7 | 9 | 5.7 | 8.8 | 5 | 5 | 4 | 7.7 | 3.4 | 5.4 |
| 1 | 2 | 2.1 | 0 | 1 | 3 | 2 | 2 | 1.5 | 1 | 1.5 | 0.8 | 0 | 2 | 1 | 1.3 | 0.6 | 2 |
| 6 | 3 | 1.3 | 9 | 5 | 4 | 3 | 1 | 2.4 | 5 | 2.4 | 4.5 | 4 | 3 | 6 | 5.4 | 4.5 | 2.8 |
| 6 | 6 | 5.7 | 13 | 6 | 7 | 6 | 6 | 5.4 | 7 | 5.4 | 7.3 | 8 | 6 | 5 | 7.7 | 5.6 | 5.5 |
| 6 | 6 | 6.6 | 10 | 3 | 10 | 6 | 7 | 6.9 | 8 | 6.9 | 8.3 | 9 | 6 | 3 | 11 | 4.5 | 6.1 |
| 2 | 3 | 2.8 | 1 | 2 | 2 | 3 | 3 | 1.2 | 2 | 1.2 | 2.2 | 0 | 3 | 2 | 1.3 | 1.7 | 3 |
| 8 | 8 | 7.4 | 3 | 7 | 5 | 8 | 7 | 6 | 8 | 6.9 | 8.1 | 7 | 8 | 6 | 4.4 | 5.6 | 7.6 |
| 12 | 11 | 11 | 4 | 12 | 5 | 11 | 11 | 10 | 8 | 10 | 7.5 | 10 | 11 | 13 | 5.7 | 12 | 11 |
| 5 | 5 | 4.3 | 2 | 5 | 5 | 5 | 4 | 5.7 | 8 | 5.1 | 7.9 | 4 | 5 | 6 | 4.4 | 4.5 | 5.4 |
| 2 | 2 | 2.1 | 1 | 7 | 1 | 2 | 2 | 2.7 | 1 | 2.7 | 0.6 | 3 | 2 | 7 | 0.3 | 6.2 | 2.4 |
| 2 | 4 | 3.9 | 1 | 1 | 1 | 4 | 4 | 4.5 | 1 | 4.5 | 1.6 | 3 | 4 | 2 | 1 | 1.1 | 3.9 |
| 3 | 4 | 5.9 | 9 | 6 | 6 | 4 | 6 | 5.1 | 5 | 5.7 | 5.3 | 5 | 4 | 6 | 5.7 | 6.8 | 4 |
| 7 | 8 | 6.9 | 9 | 11 | 9 | 8 | 7 | 5.7 | 13 | 5.4 | 1.3 | 8 | 8 | 10 | 13 | 11 | 8.3 |
| 7 | 6 | 6.4 | 10 | 9 | 3 | 6 | 6 | 5.7 | 4 | 5.4 | 4.5 | 5 | 6 | 9 | 5.4 | 9.6 | 5.9 |
| 0 | 1 | 1.1 | 0 | 2 | 1 | 1 | 1 | 1.2 | 0 | 1.2 | 0.6 | 1 | 1 | 2 | 0.7 | 1.7 | 1.2 |
| 1 | 5 | 3.4 | 0 | 2 | 3 | 5 | 3 | 3.9 | 2 | 3.9 | 1.6 | 2 | 5 | 2 | 2 | 1.7 | 4.6 |
| 10 | 6 | 7.7 | 3 | 2 | 7 | 6 | 8 | 9.6 | 6 | 8.7 | 4.9 | 6 | 6 | 3 | 7 | 3.4 | 6 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

The half life of a protein is the time it takes before only half of the protein pool for that particular protein is left. The half life of proteins is highly dependent on the presence of the N-terminal amino acid, thus

overall protein stability. Half life of all the proteins selected for the study showed a value $\geq$ 30, except fusion protein and nucleo capsid protein (Table 3). At neutral pH, the fraction of negatively charged

**Table3: Physico-chemical characteristics of Proteins of PD ruminants predicted by ProtParam**

| Proteins | No.of AA | Molwt (kDa) | pI | Q | EC | Half Life (in Hours) | II (%) | AI | GRAVY |
|---|---|---|---|---|---|---|---|---|---|
| Fusion Protein | 149 | 16305 | 9.72 | +ve | 5960 | 2.8 | 54.02 | 118.46 | -0.102 |
| F-Protein | 125 | 13440.6 | 9.36 | +ve | 1490 | 20 | 52.96 | 116.24 | -0.162 |
| NucleoProtein | 116 | 12698.8 | 5.16 | -ve | Nil | 1.9 | 51.97 | 54.83 | -0.093 |
| L-protein | 2183 | 247357.6 | 7.54 | +ve | 293595 | 30 | 37.65 | 94.97 | -0.150 |
| H-Protein | 609 | 68820.4 | 6.79 | -Ve | 70540 | 30 | 44.62 | 98.52 | -0.100 |
| Nucleo Capsi Protein | 94 | 10340.2 | 5.90 | -ve | Nil | 20 | 44.70 | 44.68 | -1.400 |
| C-protein | 177 | 20153.5 | 9.69 | +ve | 22585 | 30 | 55.64 | 84.92 | -0.331 |
| V-Protein | 298 | 32422.6 | 4.78 | -ve | 33920 | 30 | 49.69 | 66.38 | -0.754 |
| Large Protein | 2183 | 247419.3 | 7.74 | +ve | 290970 | 30 | 37.88 | 94.97 | -0.154 |
| Haeagglutinin | 609 | 68820.4 | 6.79 | +ve | 69790 | 30 | 44.62 | 98.52 | -0.100 |
| Matrix protein | 335 | 37997.9 | 8.95 | +ve | 41620 | 30 | 47.91 | 95.04 | -0.145 |
| Phospho Protein | 509 | 54859.7 | 5.02 | -ve | 26025 | 30 | 48.50 | 81.04 | -0.635 |
| M-Protein | 335 | 38059.1 | 8.84 | +ve | 41620 | 30 | 49.78 | 97.10 | -0.126 |
| P-Protein | 509 | 54966.9 | 5.03 | -ve | 28670 | 30 | 51.24 | 80.28 | -0.644 |
| N-Protein | 525 | 57687.5 | 5.20 | -ve | 43890 | 30 | 46.56 | 90.90 | -0.266 |
| RNA dependent RNA polymerase | 2183 | 247434.4 | 7.63 | +ve | 299220 | 30 | 36.55 | 94.93 | -0.146 |
| Non Structural Protein | 177 | 20278.8 | 9.93 | +ve | 26470 | 30 | 52.25 | 89.83 | -0.246 |
| Hypothetic Protein I | 398 | 31337.9 | 4.58 | -ve | 20190 | 30 | 54.04 | 6.07 | -0.725 |
| Hypothetic Protein II | 177 | 1993.1 | 9.92 | +ve | 20970 | 30 | 56.28 | 83.79 | -0.354 |
| Polymerase | 2123 | 240497.5 | 7.82 | +ve | 291645 | 30 | 36.38 | 94.23 | -0.16 |

*AA= amino acid; pI= Isoelectric point; Q= Net Charge; II= Instability index; GRAVY=Grand average of hydropathicity*

**Table 4: SOPMA Prediction Result for Secondary Structure**

| *Protein* | Alpha helix % | Extended strand % | Beta turn % | Random coil % |
|---|---|---|---|---|
| Fusion Protein | 78.52 | 4.03 | 2.68 | 14.77 |
| F-Protein | 84 | 4 | 00 | 12.80 |
| NucleoProtein | 33.62 | 6 | 00 | 61.21 |
| L-protein | 49.56 | 11.73 | 3.80 | 34.91 |
| H-Protein | 26.27 | 27.09 | 4.76 | 41.87 |
| Nucleo Capsi Protein | 10.64 | 8.51 | 0.0 | 80.85 |
| C-protein | 49.15 | 8.47 | 2.82 | 39.55 |
| V-Protein | 18.79 | 16.44 | 7.38 | 57.38 |
| Large Protein | 49.56 | 11.73 | 3.80 | 34.91 |
| Haeagglutinin | 26.27 | 27.09 | 4.76 | 41.87 |
| Matrix protein | 21.19 | 27.16 | 5.97 | 45.67 |
| Phospho Protein | 32.42 | 8.25 | 6.48 | 52.85 |
| M-Protein | 22.39 | 26.57 | 7.16 | 43.88 |
| P-Protein | 32.61 | 7.86 | 6.09 | 53.44 |
| N-Protein | 48.38 | 8.19 | 3.62 | 39.81 |
| RNA dependent RNA polymerase | 49.93 | 11.73 | 3.76 | 34.59 |
| Non Structural Protein | 49.72 | 7.91 | 3.39 | 38.98 |
| Hypothetic Protein I | 23.15 | 9.40 | 5.37 | 62.08 |
| Hypothetic Protein II | 45.20 | 11.86 | 3.95 | 38.98 |
| Polymerase | 49.98 | 12.25 | 4.24 | 33.54 |

**Parenthesis shows Number of secondary structure repeats**
**Window width 17; similarity threshold 8; Number of states 4**

**Table 5: SVMprot predicted protein families**

| Protein | Predicted function with P values | P-value | R-value |
|---|---|---|---|
| Fusion Protein | Manganese-binding | 1.3 | 68.5 |
| | Iron-binding | 1.2 | 65.4 |
| F-Protein | EC 2.4.-.-: Transferases - Glycosyltransferases | 1.7 | 78.4 |
| | Iron-binding | 1.1 | 62.2 |
| | Manganese-binding | 1.1 | 62.2 |
| Nucleoprotein | All DNA-binding | 1.0 | 58.6 |
| | Repressor | 1.0 | 58.6 |
| | mRNA-binding Proteins | 1.0 | 58.6 |
| L protein | mRNA capping | 6.7 | 99.1 |
| | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 4.4 | 98.2 |
| | Transmembrane | 1.9 | 82.2 |
| H protein | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 5.4 | 98.9 |
| | Envelope protein | 3.4 | 96.1 |
| | EC 4.1.-.-: Lyases - Carbon-Carbon Lyases | 2.0 | 83.9 |
| nucleocapsid protein | Magnesium-binding | 1.0 | 58.6 |
| C protein | Zinc-binding | 1.2 | 65.4 |
| | Nuclear Receptors | 1.2 | 65.4 |
| V protein | Zinc-binding | 1.9 | 82.2 |
| | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 1.6 | 76.2 |
| | Cell adhesion | 1.2 | 65.4 |
| Large protein | mRNA capping | 6.6 | 99.1 |
| | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 4.6 | 98.4 |
| | Transmembrane | 1.9 | 82.2 |
| **hemagglutinin** | All DNA-binding | 1.3 | 68.5 |
| | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 5.4 | 98.9 |
| | Envelope protein | 3.4 | 96.1 |
| | EC 4.1.-.-: Lyases - Carbon-Carbon Lyases | 2.0 | 83.9 |
| **matrix protein** | Structural protein (Matrix protein,Core protein,Viral occlusion body,Keratin) | 4.8 | 98.6 |
| **Phospho protein** | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 5.1 | 98.8 |
| | Zinc-binding | 1.2 | 65.4 |
| **M protein** | Structural protein (Matrix protein,Core protein,Viral occlusion body,Keratin) | 5.9 | 99.0 |
| **P protein** | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 4.4 | 98.2 |
| | Zinc-binding | 1.1 | 62.2 |
| **N protein** | Transmembrane | 3.2 | 95.2 |
| | All lipid-binding proteins | 2.8 | 92.9 |
| | Zinc-binding | 1.6 | 76.2 |
| **RNA-dependent RNA polymerase** | Transmembrane | 3.2 | 95.2 |
| | All lipid-binding proteins | 2.8 | 92.9 |
| | Zinc-binding | 1.6 | 76.2 |
| **nonstructural protein** | Zinc-binding | 1.2 | 65.4 |
| | Metal-binding | 1.0 | 58.6 |
| **Hypothetical protein-I** | Zinc-binding | 1.2 | 65.4 |
| | Metal-binding | 1.0 | 58.6 |
| **Hypothetical protein-II** | Zinc-binding | 1.1 | 62.2 |
| | Metal-binding | 1.0 | 58.6 |
| **polymerase** | mRNA capping | 5.5 | 98.9 |
| | EC 2.7.-.-: Transferases - Transferring Phosphorus-Containing Groups | 3.0 | 94.2 |
| | Transmembrane | 1.8 | 80.4 |

residues implies information about the location of the protein. Intracellular proteins tend to have a higher fraction of negatively charged residues than extra cellular proteins. At neutral pH, nuclear proteins have high relative percentage of positively charged amino acids. Nuclear proteins often bind to the negatively charged DNA, which may regulate gene expression or help to fold the DNA. Nuclear proteins often have a low percentage of aromatic residues (Andrade *et al*., 1998).The nucleocapsid protein, nuclear protein and N-protein have very low or zero percentage of aromatic amino acids (Table 2) .The aliphatic index of a protein is a measure of the relative volume occupied by aliphatic side chain of the following amino acids; alanine, valine, leucine and isoleucine. An increase in the aliphatic index increases the thermo stability of globular proteins. The fusion protein has more stability than other proteins as its AI (>100) is more than other proteins.

**Table 6: Transmembrane Protein Predicted by Sosui Server**

| Protein | N Terminal | Transmembrane Sequence | C Terminal | Type | Length |
|---|---|---|---|---|---|
| Fusion protein | 38 | FAGAVLAGVALGVATAAQITAGV | **60** | **Primary** | **23** |
| Haemagglutinin | 31 | VIERPYILLGVLLVMFLSLIGLL | **53** | **Primary** | **23** |

**Table 7: Characters of Model proposed by Swiss-model**

| Protein | Residues | Template | Seq.Identity (%) | E-value |
|---|---|---|---|---|
| **Fusion Protein** | 1 to 149 | 2b9bB (1.70 Å) | 35.57 | 0.00e-1 |
| **F-Protein** | 1 to 125 | 2b9bB (2.85 Å) | 35.2 | 0.00e-1 |
| **L-protein** | 1760 to 1965 | 2plwA (1.70 Å) | 17.062 | 4.10e-8 |
| **H-Protein** | 188 to 609 | 3alxB (3.15 Å) | 41.706 | 0.00e-1 |
| **V-Protein** | 211 to 283 | 2hyeB (3.10 Å) | 39.726 | 8.30e-29 |
| **Large Protein** | 1760 to 1965 | 2plwA (1.70 Å) | 17.062 | 4.30e-8 |
| **Haeagglutinin** | 188 to 609 | 3alxB (3.15 Å) | 41.706 | 0.00e-1 |
| **Phospho Protein** | 459 to 508 | 1oksA (1.80 Å) | 60 | 2.53e-10 |
| **P-Protein** | 459 to 509 | 1oksA (1.80 Å) | 56.863 | 2.00e-15 |
| **RNA dependent RNA polymerase** | 1760 to 1963 | 2plwA (1.70 Å) | 17.062 | 3.20e-7 |
| **Polymerase** | 1700 to 1903 | 2plwA (1.70 Å) | 16.588 | 1.70e-7 |

The secondary structure predicted values (Table 4).The fusion protein showed highest alpha helix (>75%) and lowest for nucleo capsid protein (<10). The beta turns prediction showed lowest in nuclear protein and highest in structural proteins. The functional families predicted for various proteins (Table 5) describes the hypothetical proteins as non structural one wsith metal binding properties. The fusion protein is belongs to manganese binding one and large proteins, N-proteins, polymerase are transmembrane one. The transmembarne property is not reflected in transmembarne prediction by SOSUI server, hence the SVM prediction of transmembrane property of the large proteins, N-proteins, polymerase may be due to artifacts or due to viral origin. The polymerases are classified as transferase. The SOSUI server predicted (Table 6) the fusion protein and haemagglutinin has transmembarne sequences. Since the fusion protein and haemagglutinin protein helps in the entry of the virus in the host cell transmembarne sequence is essentially necessary fro its function. The 3D model proposed are enlisted (Table 6). The sequence identity and e-values are in the acceptable limit. Further the 3D structures (Figure1) are validated by using VADAR server. All the 3D models are good concerned with various parameters like secondary structure, hydrogen bonds, dihedral angles, accessible surface area and packing volume. Ramachandran plot (Figure 2) also depicts the validity of the structure.

**Table 8: VADAR validation DATA of the predicted 3D model of the proteins**

| Secondary Structure | | | | | | |
|---|---|---|---|---|---|---|
| Statistic / PDB ID | Observed | | | | | Expected |
| | 2b9bB | 2plwA | 3alxB | 2hyeB | 1oksA | |
| Helix | 35% | 34% | 3% | 11% | 79% | - |
| Beta | 30% | 31% | 58% | 32% | 0% | - |
| Coil | 34% | 34% | 37% | 55% | 20% | - |
| Turn | 16% | 15% | 9% | 8% | 14% | - |
| **Hydrogen Bonds** | | | | | | |
| Mean HBond distance | 2.2 | 2.1 | 2.2 | 2.3 | 2.2 | 2.2. |
| Mean Hbond Energy | -1.7 | -2.0 | -1.8 | -1.5 | -1.7 | -2.0 |
| Re with Hbonds | 75% | 74% | 63% | 52% | 90% | 75% |
| **Dihedral Angles** | | | | | | |
| Mean Helix Phi | -64.3 | -64.4 | -61.5 | -57.7 | -65.1 | -65.3 |
| Mean helix Psi | -40.2 | -39.0 | -39.0 | -46.4 | -38.6 | -39.4 |
| Res with Gauche+ Chi | 55% | 59% | 41% | 53% | 60% | 55% |
| Res with Gauche- Chi | 12% | 9% | 19% | 19% | 10% | 20% |
| Res with Trans Chi | 31% | 31% | 38% | 26% | 29% | 25% |
| Mean Chi Gauche+ | -63.6 | -66.0 | -63.5 | -65.6 | -71.4 | -66.7 |
| Mean Chi Gauche- | 57.8 | 65.4 | 60.6 | 54.7 | 69.1 | 64.1 |
| Mean Chi trans | 167.2 | 169.8 | 166.9 | 165.2 | 169.5 | 168.6 |
| StiLdev of Chiu pooled | 13.32 | 11.79 | 14.42 | 15.37 | 13.27 | 15.70 |
| Mean Omega(\| omega \| >90) | 79.8 | 178.9 | 179.8 | -179.9 | 178.9 | 180.0 |
| Res with (\| omega \| >90 | 0% | 0% | 0% | 0% | 0% | - |
| **Accessible surface area (ASA)** | | | | | | |
| Total ASA | 26443.4 | 9251.4 | 24051.9 | 10687.4 | 4217.4 | 17253.8 |
| ASA of Backbone | 2708.8 | 1154.1 | 3130.1 | 1354.9 | 254.0 | - |
| ASA of side chain | 23734.6 | 8097.3 | 20921.9 | 9332.5 | 3963.5 | - |
| ASA of C | 17400.5 | 5430.9 | 14691.3 | 6954.9 | 2626.4 | - |
| ASA of N | 2015.7 | 744.0 | 1963.1 | 691.8 | 390.5 | - |
| ASA of N+ | 677.4 | 781.7 | 1399.2 | 530.7 | 443.6 | - |
| ASA of O | 5439.3 | 1903.8 | 5040.6 | 1955.4 | 455.4 | - |
| ASA of O- | 798.3 | 357.5 | 864.5 | 510.4 | 224.1 | - |
| ASA of S | 112.2 | 33.5 | 93.3 | 44.1 | 77.4 | - |
| Exposed Nonpolar ASA | 17171.7 | 5356.5 | 14369.3 | 6872.5 | 2462.0 | 16130.5 |
| Exposed Polar ASA | 6646.1 | 2252.8 | 5662.8 | 2147.0 | 517.9 | 5288.7 |
| Exposed Charged ASA | 2626.4 | 1642.0 | 4019.8 | 1667.9 | 1237.5 | 5024.2 |
| Side exposed nonpolar ASA | 17136.9 | 5298.0 | 14297.1 | 6845.1 | 2461.3 | - |
| Side exposed polar ASA | 4031.1 | 1233.2 | 2678.5 | 852.5 | 292.8 | - |
| Side exposed Polar ASA | 2566.5 | 1566.0 | 3946.2 | 1634.8 | 1209.4 | - |
| Fractuion Nonpolar ASA | 0.65 | 0.58 | 0.60 | 0.64 | 0.58 | 0.61 |
| Fraction Polar ASA | 0.25 | 0.24 | 0.24 | 0.20 | 0.12 | 0.20 |
| Fraction charged ASA | 0.10 | 0.18 | 0.17 | 0.16 | 0.29 | 0.19 |
| Mean residue ASA | 54.9 | 50.8 | 45.7 | 59.4 | 78.1 | - |
| Meran Fract ASA | 0.3 | 0.3 | 0.3 | 0.3 | 0.4 | - |
| % side chain ASA hydrophobic | 37.81 | 23.53 | 30.03 | 34.66 | 20.46 | - |
| **VOLUME** | | | | | | |
| Total volume (Packing) | 61088.6 | 2581.6 | 76182.7 | 23934.1 | 7095.6 | 62938. |
| Mean residue volume | 126.7 | 141.8 | 144.8 | 133.0 | 131.4 | 125.0 |
| Mean Fract volume | 1.0 | 1.0 | 1.1 | 1.0 | 0.9 | 1.0 |
| Molecular weight | 51820.98 | 21-82.46 | 58933.04 | 19585.88 | 6418.75 | |

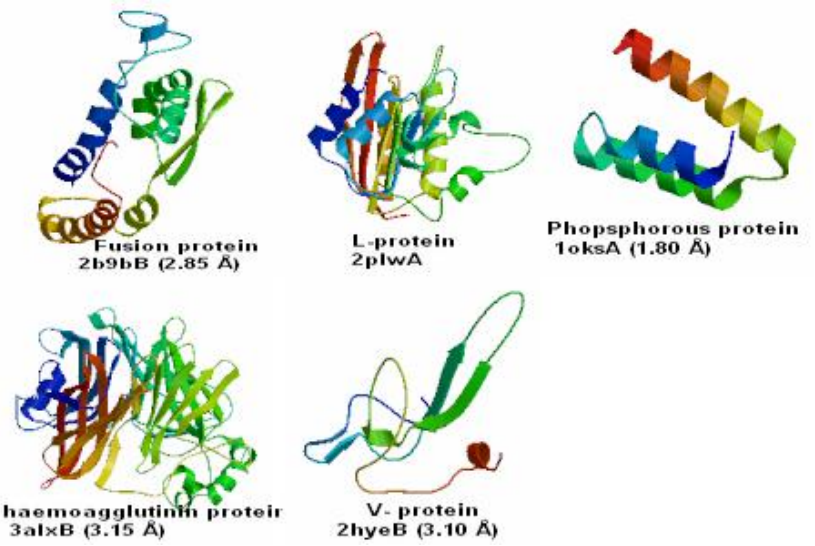


**Figure 1: PDB structure of Various Proteins from PPR**

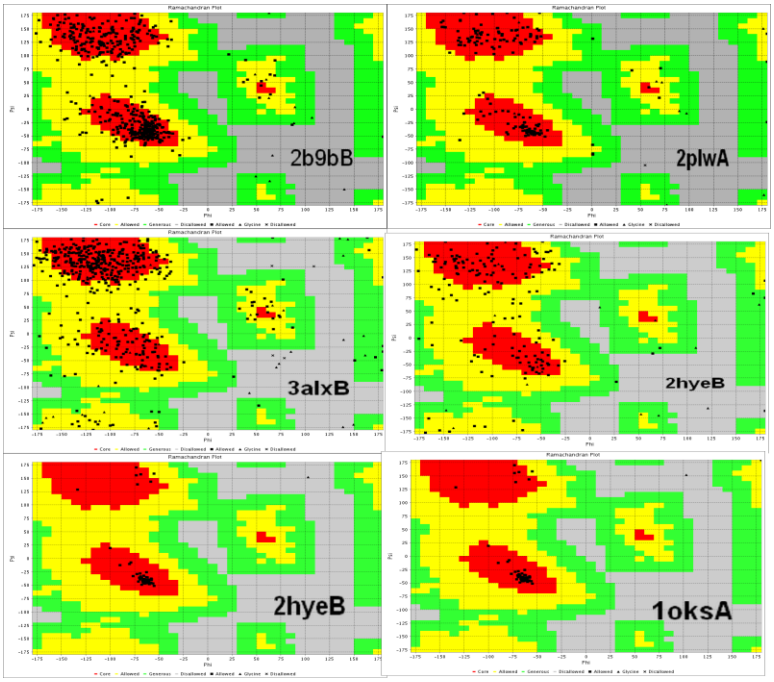


**Figure 2: Ramchandran Plot of Selected Protein Sequences**

Thus in this study we observed that out of twenty proteins selected five proteins are glycine rich and others are leucine rich one. The purification procedure for all the proteins except fusion protein and nucleo capsid protein require casual absorbance index procedure. As the fusion protein and nucleo capsid protein are absence of tryptophan a sophisticated purification method is required successful extraction. The proteins of PPR appear better stability, an indication of resistance to mutation. A further study on mutation considering this view is needed. The stability of the protein is also obvious from the secondary structure study, 3D model generated and its validation data including Ramachandran plot. We observed only two proteins of the PPR have Trans membrane moiety, a further study using more sophisticated method is needed to substantiate this finding.

**REFERENCES**

**Abraham G, Sintayehu A, Libeau G, Albina E, Roger F and Laekemariam Y(2005).** Antibody seroprevalences against peste des petits ruminants (PPR) virus in camels, cattle, goats, and sheep in Ethiopia. *Preview of Veterinary Medicine* **70** 51–57.

**Akakpo AJ, Deconinck P, Amegatse K, Kaboret Y and Oudar J (1996).** Une épidémie de la Peste des petits ruminants (PPR) en élevage périurbain à Dakar: importance épidémiologique et médicale. *Review of Medicine Veterinary* **147** 447–452.

**Andrade MA, O'Donoghue SI and Rost B (1998).** Adaptation of protein surfaces to subcellular location. *Journal of Molecular Biology* **276**(2) 517-525.

**Arnold K, Bordoli L, Kopp J and Schwede T (2006).** The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **22** 195-201.

**Barrett T, Visser IKG, Mamaev, Van Bressem MFand Osterhaus ADME (1993).** Dolphin and porpoise morbilliviruses are distinct from phocid distemper virus. *Virology* 1010– 1012.

**Betts MJ and Russell RB (2003).** Amino acid properties and consequences of subsitutions. In *Bioinformatics for geneticists* edition MR Barnes and IC Gray Chichester: Wiley.

**Devireddy LR, Raghavan R, Ramachandran S and Shaila MS (1999).** The fusion protein of peste des petits ruminants virus is a hemolysin. *Archeology of Virology* **144**(6) 1241-1247.

**Gargadennec L and Lalanne A (1942).** La peste des petits ruminants *Bulletin of Survey of Zoological Experiment* **5**(1) 16- 21.

**Gasteiger E (2005).** (In) John M Walker (edition): The Proteomics Protocols Handbook Humana Press 571-607.

**Guex N and Peitsch MC (1997).** SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18** 2714-2723.

**Hamdy FM, Dardiri AH, Nduaka O, Breese SS and Ihemelandu EC (1976)**. Etiology of the stomatitis-pneumoenteritis complex in Nigerian dwarf goats. *Canadian Journal of Comparative Medicine* **40** 276-284.

**Iemelandu EC, Naduaco O and Ojukwa EM (1985).** Hyperimmune serum in the control of Peste des petites ruminnats. *Tropical Animal Health Production* **17** 83-88.

**Jones L, Giavedoni L, Saliki JT, Brown C anb Mebus C (1993).** Protection of goats against peste des petits ruminants with a vaccinia virus double recombinant expressing the F and H genes of rinderpest virus. *Vaccine* **11** 961–964.

**Mitaku S, Hirokawa T and Tsuji T (2002).** Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. *Bioinformatics* **18** 608–616.

**Nanda YP, Chatterjee AK, PurohitA, Diallo A and Innui K (1996).** The isolation of peste des petits ruminants virus from Northern India. *Veterinary Microbiology* **51** 207-216.

**Pegram RG and Tereke F (1981).** Observation on the health of Afar livestock. *Ethiopian Veterinary Journal* **5** 11-14.

**Schwede T, Kopp J, Guex N and Peitsch MC (2003).** SWISS-MODEL: an automated protein homology-modeling serv. *Nucleic Acids Research* **31** 3381-3385.

**Scott GR (1981).** Rinderpest and peste des petits ruminants. In: Gibbs EPJ edition Virus Diseases of Food Animals **2** Academic Press London UK 401-432.

**Shaila MS, Shamki D, Morag A, Forsyth A and Diallo A (1996).** Geographical distribution and epidemiology of peste des petits ruminants viruses. *Virus Research* **43** 149-153.

**Taylor WP (1979).** Protection of goats against peste-des-petits ruminants with attenuated rinderpest virus. *Research in Veterinary Science* **27** 321-324.

**Wakwaya K (1997).** Epidemiological and serological investigations on ovine respiratory diseases and observation on vaccinated sheep on the Highland of North Shoa/Sheno. In *Doctor in Veterinary Medicine Thesis*. Debre Zeit Faculty of Veterinary Medicine.

**Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RFSykes BD and Wishart DS (2003).** VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Researches* **31(13)** 3316-3319.

Y**ayehrad TF (1997).** Epidemiological survey of Peste des petits ruminants and Contagious pleuropneumonia in selected areas of Ethiopia. In *Doctor in Veterinary Medicine thesis*. Debre Zeit Faculty of Veterinary Medicine.