

Research Article

Structural Insight of *Bos Indicus* ITGB6 Receptor Gene for FMDV Host Tropism

RANI SINGH, RAJIB DEB, UMESH SINGH, T. V. RAJA, RANI ALEX, SUSHIL KUMAR, RAFEEQUE RAHMAN ALETHODI, JAGJIT SAHU², RUPUM DATTA¹, GYANENDRA SENGAR, SHEETAL SHARMA

Molecular Genetics Laboratory, Indian Council of Agricultural Research–Central Institute for Research on Cattle, Meerut-250 001, Uttar Pradesh, India; ¹ Division of Veterinary Biotechnology, Indian Veterinary Research Institute, Izatnagar, Uttar Pradesh, India; ²Department of Life Science and Bioinformatics, Assam University, Silchar-11, Assam, India.

Abstract | Integrin beta 6 (ITGB6) is a component of FMD virus host tropism receptor in cattle. In the present study, we aimed to identify the structural variation of ITGB6 receptor gene of *Bos indicus* and their comparative differences with *Bos taurus* sequence. The study revealed that, in spite of some pivotal differences, the taurine and zebu share all common structural elements of the ITGB6 molecules. Further, study pertaining to SNP (Single nucleotide polymorphism) identification and functional genomics are ongoing to find out the genetic aspect of ITGB6 for conferring host resistance in cattle.

Keywords | ITGB6, FMD, *Bos indicus*, *Bos taurus*, Structure, AHI48847, ABH04286

Editor | Kuldeep Dhama, Indian Veterinary Research Institute, Uttar Pradesh, India.

Received | October 18, 2014; **Revised** | October 19, 2014; **Accepted** | October 20, 2014; **Published** | November 05, 2014

***Correspondence** | Rajib Deb, Central Institute for Research on Cattle, Meerut, India; **Email**: drrajibdeb@gmail.com

Citation | Singh R, Deb R, Singh U, Raja TV, Alex R, Kumar S, Alethodi RR, Sahu J, Datta R, Sengar G, Sharma S (2014). Structural insight of *Bos indicus* ITGB6 receptor gene for FMDV host tropism. *Adv. Anim. Vet. Sci.* 2 (10): 549-556.

DOI | <http://dx.doi.org/10.14737/journal.aavs/2014/2.10.549.556>

ISSN (Online) | 2307-8316; **ISSN (Print)** | 2309-3331

Copyright © 2014 Singh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Integrins are one of the important biologically active proteins responsible for FMD virus host interaction. Functional integrins consist of two noncovalently bound transmembrane glycoprotein subunits viz. alpha and beta. Each subunit is composed of a large extracellular domain of a and b residues (120–180 and 90–110 kDa, respectively), a single transmembrane domain and a small, C-terminal cytoplasmic domain (Stewart and Nemerow, 2007). In mammals, the integrin family consists of around 24 different heterodimers, each of which has a distinct tissue distribution. FMDV has been shown to use four kinds of integrins viz. avb1, avb3, avb6 and avb8 as receptors to initiate infection. Virus attachment to the integrin receptors is mediated by an arginine–glycine–aspartic acid (RGD) tripeptide located on VP1 loop of FMDV. The integrin receptors of FMDV have been inves-

tigated extensively in cell culture which have shown that avb6 but not avb3 are expressed constitutively on the epithelial cell surfaces at the site of viral replication to allow initiation of infection (Berinstein et al., 1995; Jackson et al., 2000, 2002, 2004).

In the present study we aimed to analyse the structural prediction of *Bos indicus* Integrin beta 6 (ITGB6) receptor gene and comparative evaluation with *Bos taurus*.

MATERIALS AND METHODS

SEQUENCE INFORMATION

We have obtained the full length Coding Sequence (CDS) of *Bos indicus* ITGB6 from GenBank with the accession number AHI48847. The complete CDS of zebu ITGB6 was 2367 base pair in length with 788 amino acid residues. *Bos taurus* reference sequences

included in the analysis was taken from GenBank accession number ABH04286.

PRIMARY AND SECONDARY STRUCTURE ANALYSIS

To get into the chemistry of the sequences for structural analysis, the primary structure information of the protein sequences for Accession nos. ABH04286 and AHI48847 was studied using ProtParam tool (<http://expasy.org/cgi-bin/protparam>) (Gasteiger et al., 2005) at the ExPasy Proteomic Server. Various physico-chemical parameters such as molecular weight, isoelectric point, instability index, aliphatic index, and grand average hydropathy (GRAVY) were computed. The CONCORD (Wei et al., 2011) and PSIPRED server (Buchan et al., 2010) were used for the prediction of secondary structure information of the protein.

MODEL CONSTRUCTION AND OPTIMIZATION

Model construction using appropriate computational tools has become very common practice due to their efficiency in reliable theoretical 3-D model of protein from its primary amino acid sequence. For identification of suitable templates, DELTA-BLAST (domain enhanced lookup time accelerated BLAST) against PDB was performed. DELTA-BLAST was preferred against normal BLASTP because the retrieval accuracy and sensitivity towards protein analysis is more in case of DELTA-BLAST than normal BLASTP. Discovery Studio version 3.5 (DS3.5; Accelrys Inc. San Diego, CA, USA) was used to carry out the homology modelling. DS3.5 provided 10 different models, ranked according to their discrete optimized protein energy (DOPE) score (a statistical potential to assess the quality of the models). The model with lowest DOPE score was selected as the best model. The quality of the model for Acc oxidase in *Tinospora* was evaluated by a number of tools to test the internal consistency and reliability of the model. The Structural Analysis and Verification Server (SAVeS) (<http://nihserver.mbi.ucla.edu/SAVES/>) was used which is a metasever which incorporates PROCHECK (Laskowski et al., 1996), ERRAT (Colovos et al., 1993), VERFIY3D (Eisenberg et al., 1997) and

PROVE (Pontius et al., 1996) programs for model validation. PROCHECK analysis, which quantifies the residues in available zones of Ramachandran plot, was used to assess the stereo-chemical quality of the model.

MD SIMULATION

GROMACS 4.6 (Groningen Machine for Chemical Simulations) (Van et al., 2005) package was used to run MD simulations to assess the structural integrity of the model. To investigate the stability and dynamics of the modelled protein, we considered the GRO-MOS96 43A1 force field and the flexible TIP3P water model (Walter et al., 1999). A high performance CentOS6.0 cluster computer was used for running the MD simulation. The model was surrounded by a periodic box that extends 11 Å from the protein atoms. The protonation states of all the ionisable amino acids were determined at pH 7.0. To neutralize the system, sodium counter ions were added replacing random water molecules. The Energy minimization was performed using steepest descent algorithm for 10,000 steps. A 1-ns position restrained and a 10-ns production MD simulation was performed for each simulation system at constant pressure (1 bar) and temperature (300 K). Covalent bonds in the enzymes and water molecules were constrained using the SHAKE and SETTLE algorithms, respectively. A twin cut-off scheme of 9 Å was implemented for treating long-range and van der Waals interactions. Electrostatic interactions were computed using the particle mesh Ewald (PME) method. The time step for MD simulation was 2 fs and the snapshots were saved every 1 ps. The analysis was performed using visual molecular dynamics Grace 5 (<http://plasma-gate.weizmann.ac.il/Grace/>) programs. All computations were conducted with a high performance computer cluster.

RESULTS AND DISCUSSION

PRIMARY AND SECONDARY STRUCTURE ANALYSIS

The primary sequence information was obtained using the ProtParam tool. The lengths of both the

Table 1: The primary structure information i.e. the summary of Prot Param results

Sequence	Number of amino acids	Molecular weight	pI	Instability index	GRAVY	Aliphatic index
ABH04286	788	85954.9	5.08	43.50	-0.168	81.37
AHI48847	788	85832.8	5.05	43.72	-0.170	81.87

Table 2: The secondary structure information i.e. the summary of PSIPRED and CONCORD

Sequence	No of Coil	Percentage of Coil	No of Helix	Percentage of Helix	No of Sheet	Percentage of Sheet	Total
PSIPRED							
ABH04286	514	65.2	135	17.2	139	17.6	788
AHI48847	521	66.1	124	15.7	143	18.2	788
CONCORD							
ABH04286	502	63.7	141	17.9	145	18.4	788
AHI48847	506	64.2	138	17.5	144	18.3	788

Table 3: The significant BLAST matches to PDB proteins for the two protein sequences with accession numbers ABH04286 and AHI48847

PDB ID	Description	Query coverage (%)	Identity (%)	E value
3IJE	Chain B, Crystal Structure Of The Complete Integrin Alhavbeta3 Ectodomain Plus An Alpha/beta Transmembrane Fragment	87%	50%	0
4G1M	Chain B, Re-Refinement Of Alpha V Beta 3 Structure	86%	50%	0
4G1E	Chain B, Crystal Structure Of Integrin Alpha V Beta 3 With Coil-Coiled Tag	86%	50%	0
1JV2	Chain B, Crystal Structure Of The Extracellular Segment Of Integrin Alphavbeta3	86%	50%	0
3FCS	Chain B, Structure Of Complete Ectodomain Of Integrin Aiibb3	86%	50%	0

Table 4: SAVEs server analysis summary of the predicted model

Model validation server	Parameters	ABH04286	AHI48847
Procheck	Most favored regions (%)	87.4%	87.3%
	Additionally allowed regions (%)	11.3%	11.1%
	Generously allowed regions (%)	1.3%	1.0%
	Disallowed regions (%)	0%	0.6%
	Overall G-factor	-0.11	-0.11
Verfiy-3D	Averaged 3D-1D Score > 0.2	81.99%	fail
ERRAT	Overall Quality (%)	66.902	fail
PROVE	Prove RMS Z-score	21.616	30.614

amino acid sequences were 788 (Table 1). The molecular weight for ABH04286 and AHI48847 were 85954.9 and 85832.8. Primary sequence information also provided the theoretical pI which was 5.08 and

5.05 for ABH04286 and AHI48847 respectively. The instability index and aliphatic index were calculated to be 43.50 and 81.37 for ABH04286 whereas for AHI48847 they were 43.72 and 81.87 respectively. The tool also calculated the grand average hydropathy i.e. GRAVY which was -0.168 for ABH04286 and -0.170 for AHI48847. To insight into the secondary sequence information, 2 servers were used; CONCORD and PSIPRED (Table 2). For ABH04286, the percentage of coils, helix and sheet were found 65.2%, 17.2% and 17.6% in case of PSIPRED results whereas in CONCORD, the same were found with 63.7%, 17.9% and 18.4%. In case of AHI48847 the PSIPRED and CONCORD detected the percentage of coils, helix and sheet as 66.1%, 15.7%, 18.2% and 64.2%, 17.5%, 18.3% respectively.

MODEL CONSTRUCTION AND OPTIMIZATION

The two protein sequences were searched against PDB database and significant matches with protein sequences having 3D structure. BLAST matches for the two Accessions i.e. ABH04286 and AHI48847 were the same as these two sequences are having 3 nucleotide mismatches in between them. The matches having more than 80% query coverage were taken into account for the homology modelling. A total of 5 PDB IDs were found to have significant matches (Table 3). The PDB ID 3IJE which represents Chain B, crystal structure of the complete Integrin Alhavbeta3 Ectodomain plus an alpha/beta transmembrane fragment was found to have maximum query coverage of 87% with an identity of 50% and e-value of 0. All the 5 PDB IDs in table 4 were used as template for both of the protein sequence in homology modelling. Model building was performed after aligning the structure and sequence. DS3.5 constructed two models for the ABH04286 and AHI48847 sequences which were very homologues due to their sequence similarity as well as having the same templates.

ABH04286 MODEL

The 3D structure of the predicted model has been shown in the Figure 1-a. The model was submitted to SAVeS server for the model validation. The Procheck server provided the Ramachandran plot which presented the amino acid positions (Figure 2-a). The plot found 87.4% residues in the favourable region along with 11.3% residues in the additionally allowed and 1.3% in generously allowed regions. There was 0 residues present in the disal-

lowed region. The overall G-factor was -0.11. The Verify-3D server predicted 81.99% for the averaged 3D-1D Score > 0.2. The ERRAT and PROVE servers predicted the overall quality as 66.902 and Prove RMS Z-score to be 21.616 respectively (Table 4).

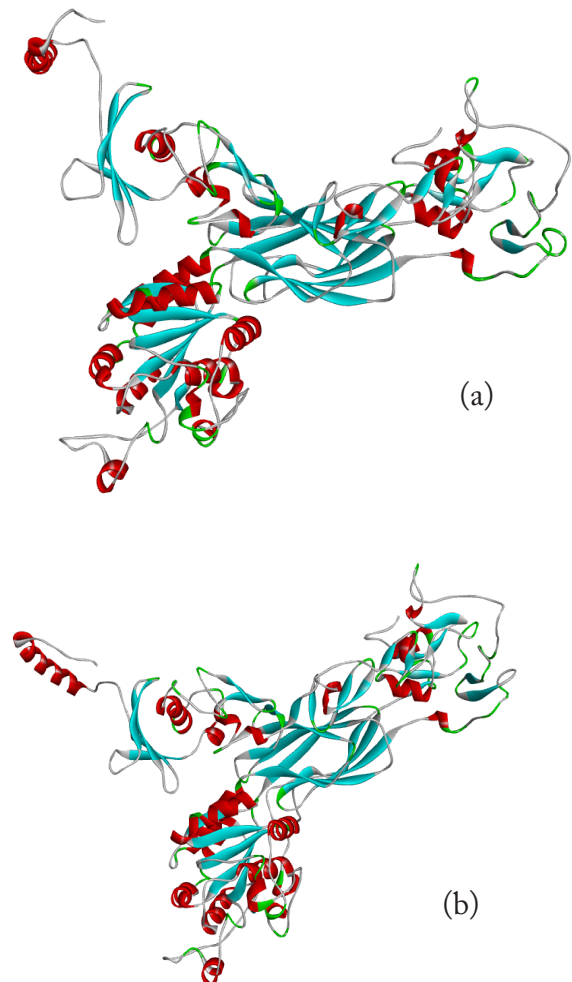


Figure 1: The predicted model through homology modelling using Discovery Studio 3.5. a) ABH04286 and b) AHI48847. The 3D structure of the predicted model has been shown here. The model was submitted to SAVeS server for the model validation.

AHI48847 MODEL

The 3D structure of the predicted model has been shown in the Figure 1-b. The Procheck server in the SAVeS server provided the Ramachandran plot which presented the amino acid positions (Figure 2-b). The plot found 87.3% residues in the favourable region along with 11.1% residues in the additionally allowed and 1% in generously allowed regions. There was 0.6% residues present in the disallowed region. The overall G-factor was -0.11 same as ABH04286.

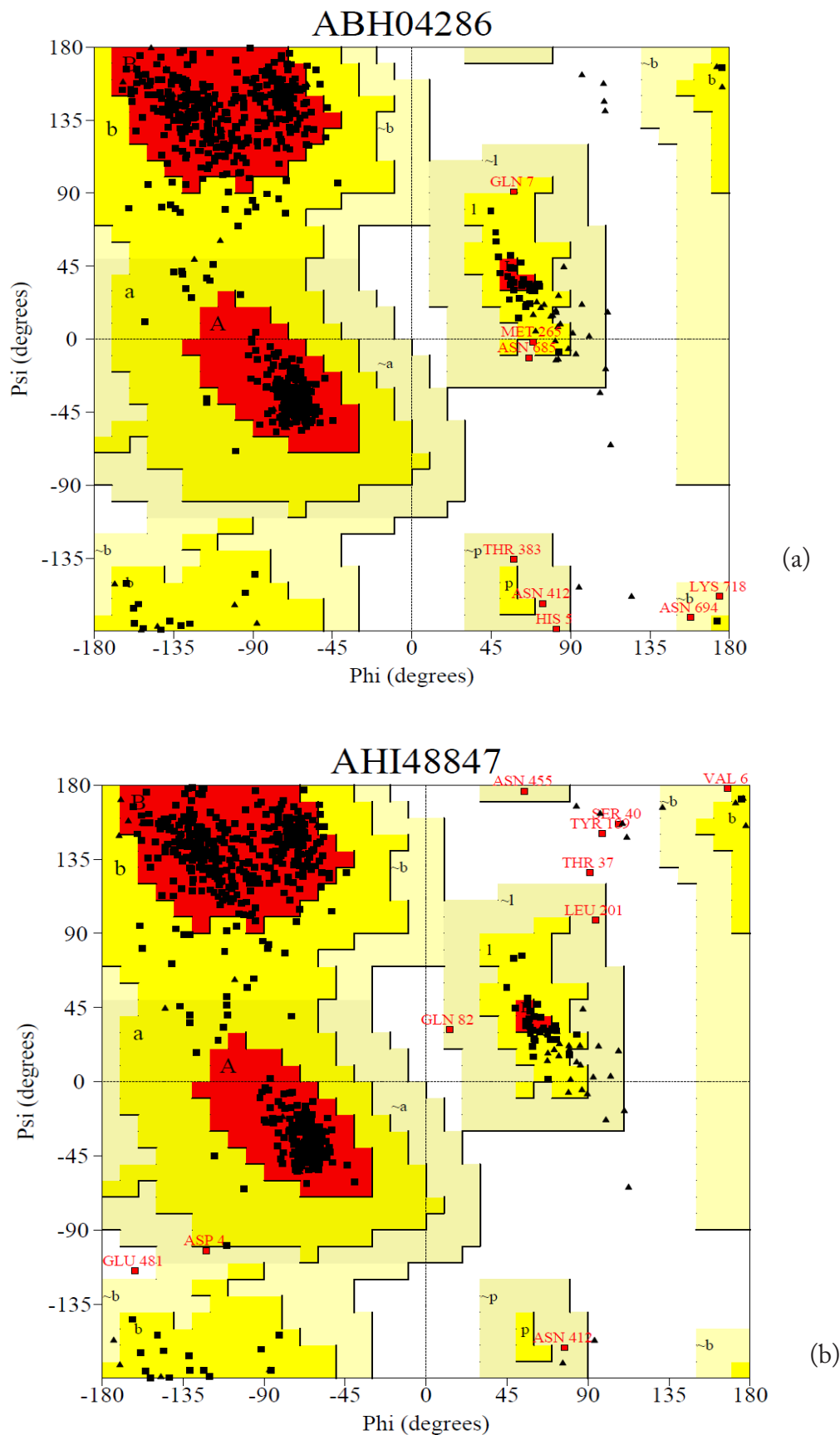


Figure 2: Ramachandran plot for the predicted model from the SAVEs server a) ABH04286 and b) AHI48847. The Procheck server provided the Ramachandran plot which presented the amino acid positions shown in the figure.

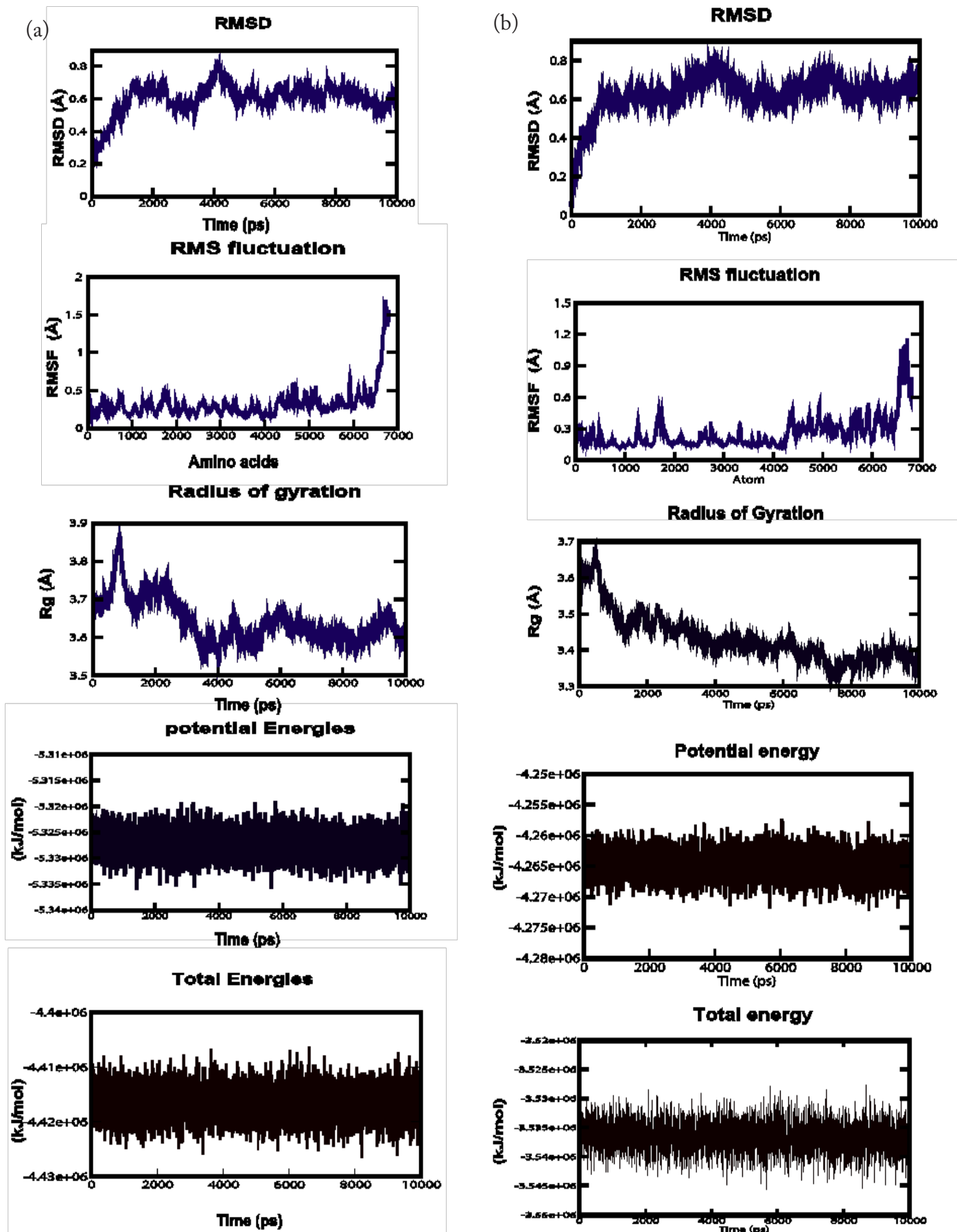


Figure 3: Root Mean Square Deviation (RMSD), Root Mean Square fluctuation (RMSf), Radius of gyration, Potential energies and Total energies plots for the modelled protein throughout the MD simulation a) ABH04286 and b) AHI48847.

The Verfy-3D server and ERRAT failed to predict the averaged 3D-1D Score > 0.2 and the overall quality. The PROVE server predicted the Prove RMS Z-score to be 30.614 (Table 4).

MD SIMULATION

ABHo₄₂₈₆ MODEL

MD simulations were performed to study the dynamic behaviour of the predicted model at 10 ns. The Root Mean Square Deviation (RMSD) graph provided the information regarding the change in structure from the original one throughout the course of simulation. The change observed in this case varied from 0.18 Å to 1 Å with an average of approximately 0.6 Å (Figure 3-a). The Root Mean Square Fluctuation (RMSF) graph represented the fluctuation of each atom among all the atoms. The graph showed a minimum at approximately 0 Å to an upper limit of 1.7 Å at the end. The gyration plot in which the radius of gyration (RG) was plotted in the Y-axis and time interval in the X-axis showed an average value of approximately 3.6 Å, the RG value ranged from around 0 Å to 3.9 Å. The potential and total energy changes throughout the simulation process were represented in the plot. The average potential energy was approximately - 5325000 KJ/mol whereas in case of total energy, it was - 4415000 KJ/mol. The Radius of gyration plot measured the compactness of the structure and the mass of the atom relative to the centre of mass of the molecule.

AHI₄₈₈₄₇ MODEL

MD simulations were performed to study the dynamic behaviour of the predicted model at 10 ns. The Root Mean Square Deviation (RMSD) graph provided the information regarding the change in structure from the original one throughout the course of simulation. The change observed in this case varied from 0 Å to 0.19 Å with an average of approximately 0.6 Å (Figure 3-b). The Root Mean Square Fluctuation (RMSF) graph represented the fluctuation of each atom among all the atoms. The graph showed a minimum at approximately 0.1 Å to an upper limit of 1.2 Å, but at the end atom, the value was 0.5 Å. The gyration plot in which the radius of gyration (RG) was plotted in the Y-axis and time interval in the X-axis showed an average value of approximately 3.5 Å, the RG value ranged from around 0 Å to 3.7 Å. The potential and total energy changes throughout the simulation process were represented in the plot. The average potential

energy was approximately - 4265000 KJ/mol whereas in case of total energy, it was - 3535000 KJ/mol. The Radius of gyration plot measured the compactness of the structure and the mass of the atom relative to the centre of mass of the molecule.

From the structural performance of ITGB6, it revealed that, despite some focal differences, the taurine and zebu cattle share all common elements of the ITGB6 molecules. Genetic makeup responsible for conferring naturally acquired FMDV resistance among zebu cattle is quite interesting and need to be initiated. At present, we are trying to identify SNPs within ITGB6 receptor gene and their association with FMDV susceptibility in cattle. Further, study on the identification of transcript variants among Zebu ITGB6 receptor gene also need to be initiated.

ACKNOWLEDGEMENT

The project (SR/WOS-A/LS-437/2012 -G) was financially supported by Department of Science and Technology, Govt. of India. The authors are thankful to Director, ICAR-CIRC, Meerut, for providing necessary facilities for conducting the study.

REFERENCES

- Berinstein A, Roivainen M, Hovi T, Mason PW, Baxt B. (1995). Antibodies to the vitronectin receptor (integrin $\alpha v \beta 3$) inhibit binding and infection of foot-and mouth disease virus to cultured cells. *J. Virol.* 69: 2664-2666.
- Buchan DW, Ward SM, Lobley AE, Nugent TC, Bryson K, Jones DT (2010). Protein annotation and modelling servers at University College London. *Nucl. Acids Res.* 38 (Web Server issue): W563-568.
- Colovos C, Yeates TO (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2(9): 1511-1519.
- Eisenberg D, Lüthy R, Bowie JU (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* 277: 396-404.
- Gasteiger E, Hoogland C, Gattiker A (2005). Protein identification and analysis tools on the ExpASY Server. In: John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press. Pp. 571-607.
- Jackson T, Sheppard D, Denyer M, Blakemore W, King AMQ (2000). The epithelial integrin $\alpha v \beta 6$ is a receptor for foot-and-mouth disease virus. *J. Virol.* 74(11): 4949-56.
- Jackson T, Mould AP, Sheppard D, King AMQ (2002). Integrin $\alpha v \beta 1$ is a receptor for foot-and-mouth disease

- virus. *J. Virol.* 76(3): 935–41.
- Jackson T, Clark S, Berryman S, Burman A, Cambier S, Mu D (2004). Integrin avb8 functions as a receptor for foot-and-mouth disease virus: role of the b-chain cytodomain in integrin-mediated infection. *J. Virol.* 78(9): 4533–40.
 - 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(4): 477–86.
 - Pontius J, Richelle J, Wodak SJ (1996). Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J. Mol. Biol.* 264(1): 121–36.
 - Stewart PL, Nemerow GR (2007). Cell integrins: commonly used receptors for diverse viral pathogens. *Trends Microbiol.* 15(11): 500–7.
 - Van DDS, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005). GROMACS: Fast, flexible, and free. *J. Comput. Chem.* 26(16): 1701–1718.
 - Walter RPS, Philippe HH, Ilario GT (1999). The GROMOS biomolecular simulation program package. *J. Phys. Chem.* 103(16): 3596–3607.
 - Wei Y, Thompson J, Floudas CA (2011). CONCORD: A consensus method for protein secondary structure prediction via Mixed Integer Linear Optimization. *Proc. R. Soc. A.* 468(2139): 831–850.