IN SILICO CHARACTERIZATION OF MX2 PROTEIN OF BUFFALO (*BUBALUS BUBALIS*)

L. Buragohain[1,*], P. Borah[1], N. Goswami[2] and R. Dutta[3]

ABSTRACT

The Myxovirus resistance protein 2 of Homo sapiens is well characterized in terms of structure and function but very little is known about myxovirus resistance protein 2 (Mx2) of Bubalus bubalis. Although a few functions of buffalo Mx2 protein are known, its structural characteristics have still remained unclear. In the present study, buffalo Mx2 protein was characterized by defining its possible structure and functions using several online and offline computational tools. Besides the physico-chemical characteristics, various post translational modifications, localization, possible domains and interactions with other probable proteins were predicted using its amino acid sequence. In addition to the secondary structure, the three-dimensional structure was predicted using I-TASSER homology modelling server and a refined model was constructed in Modeller. The refined model manifested acceptable quality and stereochemistry on validation with different programs. Furthermore, a phylogenetic analysis of the protein was done and Mx2 protein of buffalo was found to have a common ancestry with that of other bovidae family members. These curated observations might provide a sound foundation for further structural and functional characterization of buffalo Mx2 protein.

Keywords: *Bubalus bubalis*, buffalo, Mx2 protein, 3D structure

INTRODUCTION

Myxovirus resistance protein (Mx) is an interferon-induced GTP-binding protein that confers resistance to various RNA viruses infecting several vertebrates (Pavlovic *et al*., 1993). The antiviral activity of Mx protein was first reported in laboratory mice against influenza virus (Lindenmann *et al*., 1963). The Mx protein belongs to the dynamin superfamily and possesses intrinsic GTPase activity (Staeheli *et al*., 1993). The consensus properties existing in Mx proteins are tripartite GTPase domains in the N-terminal region, a dynamin signature and with a GTP effector

[1]Department of Animal Biotechnology, College of Veterinary Science, Assam Agricultural University, Assam, India, *E-mail: lukumoni54@gmail.com

[2]Bioinformatics Infrastructure Facility, College of Veterinary Science, Assam Agricultural University, Assam, India,

[3]Advanced State Biotech Hub (Assam), College of Veterinary Science, Assam Agricultural University, Assam, India

domain (GED) having a leucine zipper motif in the C-terminal region (Melen *et al*., 1992). Several isoforms of Mx protein are reported in numerous vertebrates (Horisberger and Gunst, 1991). Two Mx genes, Mx1 and Mx2 have been identified in the mouse (Staeheli and Sutcliffe, 1988). Human also possesses two Mx genes which are designated as MxA and MxB that are homologous to mouse Mx1 and Mx2, respectively (Aebi *et al*., 1989). More than one isoform of the Mx genes are documented in livestock, such as bovine (Ellinwood *et al*., 1998), buffalo (Babiker *et al*., 2016) and porcine (Muller *et al*., 1992). In bovine, the isoforms of Mx genes are Mx1 and Mx2 but Mx1 has alternative splice variants, Mx1-A and Mx1-B (Kojima *et al*., 2003).

In mouse, the Mx1 protein is localized in the cell nucleus, in contrast to Mx2 protein which is localized in the cytoplasm and both the isoforms possess antiviral activity against various RNA viruses (Staeheli *et al*., 1986; Zurcher *et al*., 1992). In human, MxA is located in the cytoplasm of cells with antiviral activity whereas MxB is found in nucleus and its antiviral activities have not yet been elucidated (Pavlovic *et al*., 1990). The Mx1-A of cattle localizes in the cytoplasm and supresses vesicular stomatitis virus (VSV) infection, but Mx1B localizes in the nucleus and its antiviral activity is yet to be documented (Nakatsu *et al*., 2004; Yamada *et al*., 2009). In case of buffalo, the antiviral activity of Mx2 protein against VSV is well documented and the gene has been characterized (Babiker *et al*., 2007; Babiker *et al*., 2016) but the protein has not yet been well characterized in relation to its structure and its subcellular localization is not clear. Besides viral infection, the Mx2 transcript and protein is also highly expressed in buffalo during early pregnancy. On the contrary, the expression pattern of Mx1 gene in pregnancy is dynamic and different experiments had different conclusions. The Mx genes are also a class of Interferon-stimulated genes (ISGs), and their levels of mRNA and protein are increased by Interferon-τ (IFN- τ) stimulation in the uterine tissue at the time of implantation in farm animals (Ott *et al*., 1998; Hicks *et al*., 2003; Spencer *et al*., 2008). Moreover, IFN-τ also stimulates expression of Mx2 gene in peripheral blood mononuclear cells during early pregnancy in bovine (Matsuyama *et al*., 2012). Although its expression is up-regulated during early pregnancy in ruminants, the exact relation is still not well understood.

The characterization of Mx2 protein of buffalo is a pre-requisite because the higher expression of Mx2 gene in early pregnancy is opening a new window for early diagnosis of pregnancy in livestock especially in cattle and buffalo. Although attempts were made to partially characterize Mx2 protein of buffalo but many facts are still to be uncovered. Therefore, the present study aimed to investigate the possible structural and functional characteristics of Mx2 protein of buffalo using various state-of-the-art bioinformatics tools.

MATERIALS AND METHODS

Sequence retrieval and physico-chemical characterization

The amino acid sequence of Mx2 protein of buffalo was retrieved in fasta format from UniPort database (http://www.uniprot.org/) having accession number A0MWD1. The basic physico-chemical properties like theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient (Gill and Von Hippel, 1989), instability index (Guruprasad *et al*., 1990), aliphatic index (Ikai, 1980) and grand

average hydropathy (GRAVY) (Kyte and Doolittle, 1982) were computed using ProtParam tool of ExPASy: SIB Bioinformatics Resource Portal (http://web.expasy.org/protparam/).

Subcellular localization, secondary structure and modifications

The subcellular localization of Mx2 protein of buffalo was predicted using Balanced Subcellular Localization Predictor (BaCelLo) (http://gpcr.biocomp.unibo.it/bacello/pred.htm) and further confirmed by online tool WoLF PSORT (https://wolfpsort.hgc.jp/). The secondary structural feature of the deduced protein was computed by the software SOPMA (Self Optimized Prediction Method with Alignment) (Geourjon and Deléage, 1995). The Intrinsic disorder regions of this protein were analysed using DISOPRED3 online program of PSIPRED-Protein Sequence Analysis Workbench (Jones *et al*., 2015; http://bioinf.cs.ucl.ac.uk/psipred/). The presence of any signal peptide and transmembrane helices was evaluated using SignalP 4.1 (Petersen *et al*., 2011; http://www.cbs.dtu.dk/services/SignalP/) and TMHMM v. 2.0 servers (Krogh *et al*., 2001; http://www.cbs.dtu.dk/services/TMHMM/), respectively. The protein was also evaluated for possession of leucine-rich nuclear export signals (NES) using NetNES 1.1 Server (Cour *et al*., 2004; http://www.cbs.dtu.dk/services/NetNES/). The online prediction servers available at the Center for Biological Sequence Analysis (CBS), Technical University of Denmark (http://www.cbs.dtu.dk/) was used to predict probable phosphorylation (NetPhos 3.1; Blom *et al*., 1999), O-linked glycosylation (NetOGlyc 4.0; Steentoft *et al*., 2013), N-linked glycosylation (NetNGlyc 1.0; http://www.cbs.dtu.dk/services/NetNGlyc/), glycation of amino groups of lysines (NetGlycate 1.0 Server; Johansen *et al*., 2006), C-mannosylation (NetCGlyc 1.0 Server; Julenius, 2007) and N-terminal acetylation (NetAcet 1.0 Server; Kiemer *et al*., 2005) sites in the putative protein.

Computational modelling of three dimensional (3D) structures

The primary amino acid sequence of buffalo Mx2 protein (A0MWD1) retrieved from UniPort database (http://www.uniprot.org/) was used for construction of three dimensional structure of the analysed protein. In the absence of experimental 3D structure of Mx2 protein of either buffalo or cattle, the online server Iterative Threading Assembly Refinement (I-TASSER; http://zhanglab.ccmb.med.umich.edu/I-TASSER/; Zhang, 2009; Roy *et al*., 2012; Yang and Zhang, 2015) was used for prediction of buffalo Mx2 protein 3D model. This program at first identified 10 best structure templates on the basis of *Z-score* from PDB library using Local Meta-Threading-Server (LOMETS) (Wu and Zhang, 2007) which contains multiple threading programs. Finally, the best five models were generated using SPICKER program (Zhang and Skolnick, 2004) and the confidence of models was evaluated on the basis of highest C-score. The cross evaluation and validation of five models were done using Structure Analysis and Verification Server (SAVES) v5 online tool (http://servicesn.mbi.ucla.edu/SAVES/). It is a meta-server for validation of protein structures and for this protein, program ERRAT (http://services.mbi.ucla.edu/ERRAT/) was used for assessment of quality factor and PROCHECK (Laskowski *et al*., 1993; Laskowski *et al*., 1996) was used for stereo-chemical evaluation. The five models were further quality checked using QMEAN server (http://swissmodel.expasy.org/qmean/cgi/index.cgi) (Benkert *et al*., 2008). On the basis of SAVES

data and QMEAN score, the best model was selected for further refinement. The refinement of the disordered region was carried out several times using Modeller 9.19 (Sali and Blundell, 1993) until the dihedral angles (φψ) of majority of the residues fell in the most favoured and allowed region of the Ramachandran plot as evaluated by PROCHECK program. Graphical inspections and representation of the models were done using UCSF Chimera (http://www.cgl.ucsf.edu/chimera/) (Pettersen *et al*., 2004). The energy profiles of the human MxB protein (PDB ID: 4WHJA; the 1st rank template used by I-TASSER) along with the refined model were evaluated by ProSA-web: protein structure analysis (Sippl, 1993; Wiederstein and Sippl, 2007).

Protein interactions and structural organization

The STRING 10.5 tool was used to analyse the interactions network of Mx2 proteins with other proteins (https://string-db.org) (Szklarczyk *et al*., 2017). The domain composition of this protein was predicted using the SMART-Simple Modular Architecture Research Tool (Letunic and Bork, 2018) (http://smart.embl-heidelberg.de).

Phylogenetic analysis

Twenty one numbers of amino acid sequences of Mx2 protein belonging to different mammalian species (cattle, buffalo, sheep, goat, horse, human etc.) and one Mx1 protein sequence of buffalo were retrieved from GenPept database (http://www.ncbi.nlm.nih.gov/).

Altogether, twenty three amino acid sequences (including the analysed protein sequence) were considered for phylogenetic analysis of buffalo Mx2 protein. The Mx1 protein sequence of buffalo (Accession No: BAG72219.1) was used as outgroup. The phylogenetic analysis was done in MEGA 7 software (Kumar *et al*., 2016). Alignment of multiple sequences was carried out using the MUSCLE program (Edgar, 2004) implemented within MEGA 7. The best-fit substitution model was estimated in MEGA 7 and selection was done on the basis of Bayesian information criteria (BIC). The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). All positions that contained alignment gaps and missing data were eliminated from the analysis. Phylogenetic tree was constructed using the Unweighted Pair Group Method with Arithmetic Mean method (UPGMA) (Sneath and Sokal, 1973) and the validity of the tree generated was tested by bootstrap analysis (Felsenstein, 1985) of 1000 pseudo-replicates. Sequence Demarcation Tool (SDT) v1.2 (Muhire, 2014) was used to estimate the pairwise identity and molecular grouping of mammalian Mx2 proteins using the same amino acid sequences that were used for construction of phylogenetic tree except the Mx1 protein sequence.

RESULTS AND DISCUSSION

Physico-chemical characterization

The ProtParam tool available at ExPASy server predicted that the full length Mx2 amino acid sequence of buffalo encoded a protein with 710 amino acid residues having a molecular weight of 81.334 kDa. The theoretical isoelectric point (pI) of this protein was found to be 6.48 with 92 negatively charged residues (Asp + Glu), while the total number of positively charged residues (Arg + Lys) was found to be 88 which were also predicted by the ProtParam tool. The Instability index was computed to be 44.57 which indicated that the protein may be slightly unstable as the value was greater than 40 (Guruprasad *et al*., 1990). The aliphatic index is used to specify the

thermal stability of a protein (Ikai, 1980), which was calculated to be 89.93 for buffalo Mx2 protein, indicating its stability at higher temperature. The Grand average of hydropathicity (GRAVY) (Kyte *et al*., 1982) of the protein was found to be -0.463, indicating the protein to be hydrophilic in nature.

Subcellular localization, secondary structure and modifications

The Mx2 protein of buffalo was predicted to be a cytoplasmic protein by the online tool Balanced Subcellular Localization Predictor (BaCelLo) (Pierleoni *et al*., 2006). Moreover, its subcellular localization was additionally confirmed to be the cytoplasm by software WoLF PSORT (Horton *et al*., 2007). The Mx2 protein of buffalo has antiviral activity against VSV (Babiker *et al*., 2007) which basically replicates in the cytoplasm of mammalian cells. This further clarifies its localization in the cytoplasm as predicted in the present study. Furthermore, the Mx2 protein of mice was documented to be a cytoplasmic protein (Staeheli *et al*., 1986; Zurcher *et al*., 1992) which also supports the current *in silico* analysis of this protein in relation to its subcellular localization. The SOPMA (Geourjon *et al*., 1995) revealed that the predominant secondary structure of this protein possessed alpha helix (48.45%) followed by random coil (27.46%), extended strand (17.89%) and beta turn (6.20%). The intrinsic disordered regions of the protein, as determined by DISOPRED3 program, revealed three prominent disordered sites and two disordered protein-protein interaction (protein binding) sites (Figure 1). Except for the first disordered and binding region which spans about initial 100 amino acids, other two prominent disordered and one protein binding sites were of small segment, thus the protein might be structurally and functionally rigid and unique.

The protein was devoid of any signal peptide, nuclear export signal and transmembrane helix as computed by SignalP 4.1 (Petersen *et al*., 2011), NetNES 1.1 (Cour *et al*., 2004) and TMHMM v 2.0 (Krogh *et al*., 2001) servers, respectively. This also suggested that the protein is not destined to be transported across the membrane and hence localized in the cytoplasm. The Mx2 protein of buffalo consisted of 24 serine, 26 threonine and four tyrosine phosphorylation sites as predicted by NetPhos 3.1 server (Blom *et al*., 1999). The NetOGlyc 4.0 server (Steentoft *et al*., 2013) predicted 16 potential O-linked glycosylation sites at the residues 2, 4, 14, 16, 17, 18, 40, 42, 51, 61, 79, 339, 343, 595, 602 and 605. This protein was predicted to have five N-linked glycosylation sites (position 59, 171, 178, 309 and 379) by the online tool NetNGlyc 1.0 but the third glycosylation site at 178 position needs further confirmatory evidences as the next amino acid after asparagine was proline in Asn-Xaa-Ser/Thr sequons (where Xaa is any amino acid except Proline) and this proline residue may preclude N-linked glycosylation in asparagine residue by rendering it inaccessible. Twenty two glycation sites at amino groups of lysine residue (9, 11, 23, 80, 266, 267, 315, 316, 390, 392, 421, 423, 449, 469, 486, 533, 572, 606, 635, 656, 689 and 691) were predicted by NetGlycate 1.0 tool (Johansen *et al*., 2006) in this protein. The protein also have single N-terminal acetylation site in the serine residue at position 2 but it lacks any C-mannosylation site as predicted by NetAcet 1.0 (Kiemer *et al*., 2005) and NetCGlyc 1.0 (Julenius, 2007) servers, respectively.

Computational modelling of three dimensional structures

Theoretical modelling of three dimensional structure of Mx2 protein of buffalo was performed

as there was insufficiency of experimental data for this considered protein. The computational modelling was executed using I-TASSER (Zhang, 2009) online server. The best five sets of three dimensional models were generated by I-TASSER having C-score ranging from -1.59 to -2.79 which was in the acceptable range (-5 to 2). Although the higher C-score value denoted higher confidence of a model and vice-versa, upon further validation and cross evaluation with the help of ERRAR and PROCHECK, the model 1 with higher C-score (-1.59) was found to be of poor quality compared to model 4 with C-score -2.41. The overall quality of model 4 determined by ERRAT was the highest (93.01%) and the stereo-chemical properties checked by PROCHECK was found to be the best among all (Table 1). The quality check in QMEAN server also additionally supported the model 4 to be the best among all five models (Table 1). According to the QMEAN global score, the highest predicted model reliability was measured that lies between 0 and 1 (1 is the highest quality of a predicted model) and for model 4, the global score was the highest (0.50). Hence, the model 4 was selected for further refinement which was done in Modeller 9.19 in several cycles. Eventually for the refined model, the quality factor estimated by ERRAT was found to be 91.45% and PROCHECK revealed all of the residues' dihedrals were within the allowed region of the Ramachandran plot. However, model 4 of I-TASSER and the refined model (Figure 2a and 2b) along with their respective Ramachandra Plot (Figure 2c and 2d) showed that there were 83.5% residues in the most favoured region and none in the disallowed region in the refined model, whereas there were 75.2% residues in the most favoured region and 0.9% residues in the disallowed region in case of model 4. Moreover, the significant increase in QMEAN score of the refined model (0.60) also suggested it to be a good quality 3D model of buffalo Mx2 protein (Table 1). The topmost (1st Rank) threading template used by I-TRASSER and the template with the highest structural similarity identified by TM-align (Zhang and Skolnick, 2005), an inbuilt feature of I-TASSER, was revealed to be experimental 3D structure of MxB protein (PDB ID: 4WHJ_A) of *Homo sapiens.* It had a TM-score of 0.79, sequence identity of 73% and query coverage of 79% with respect to human MxB protein.

Furthermore, alignment of refined model and 4WHJ_A model in TM-align produced TM-score of 0.77 (normalised by refined model) which suggested significant structural similarity because a TM-score of 0.5 to 1 generally have the same fold whereas scores below 0.3 corresponds to random structural similarity. Superimposition of the refined model and human MxB protein (Chain A of 4WHJ) produced in UCSF Chimera also revealed similar results with root mean square deviation (RMSD) of 0.44 Å (Figure 3) deciphering that both the structures were almost alike. The ProSA web-server checks the potential error in experimental and theoretical 3D models of protein and presents the overall quality as Z-score which should be within the range of the native protein structures of similar size either determined by X-ray crystallography or NMR spectroscopy. The Z-score of refined model was found to be -9.7 which was within the range of human MxB protein (-9.21), the best template (4WHJ) (Figure 4a and 4b). Moreover, the local model quality and 3D model with coloured residues in the order of increasing energy were found to be of similar pattern between the refined and template structure (Figure not shown). Thus, it indicated that the refined model could be a true reflection of an experimentally-obtained Mx2 protein structure of buffalo.

Table 1. Corresponding results of I-TASSER and refined model(s) of Mx2 protein of buffalo evaluated with ERRAT, PROCHECK and QMEAN.

Models from I-TASSER	Quality factor (%)	Ramachandran plot comparative analysis				QMEAN4 score	
		Residues in the most favoured region (%)	Residues in additionally allowed region (%)	Residues in generously allowed region (%)	Residues in disallowed region (%)	Global Score	Z-score
1	85.42	74.6	18.7	4.9	1.9	0.44	-8.74
2	85.18	74.3	20.3*	3.3	2.2	0.46	-8.35
3	89.17	73.2	18.7	5.8*	2.4	0.48	-7.67
4	93.01*	75.2	19.2	4.7	0.9	0.50	-7.22
5	82.90	73.6	18.5	5.3	2.5*	0.45	-8.56
Refined model	91.45	83.5*	15.5	0.9	0.0	0.60*	-4.38*

Superscript (*) denotes the highest value in a column.

Table 2. Top ten proteins that interact with buffalo Mx2 protein as predicted by STRING.

Protein ID	Description	Score
ISG15	Ubiquitin-like protein ISG15	0.938
RSAD2	Radical S-adenosyl methionine domain-containing protein 2	0.915
IFI44	Uncharacterized protein	0.901
USP18	Ubl carboxyl-terminal hydrolase 18	0.897
IFIH1	Uncharacterized protein	0.890
DHX58	Probable ATP-dependent RNA helicase DHX58	0.887
UBA7	Ubiquitin-like modifier-activating enzyme 7	0.876
ENSBTAG00000045588	Uncharacterized protein	0.864
RTP4	28kD interferon responsive protein	0.841
ZNFX1	NFX1-type zinc finger-containing protein 1	0.828

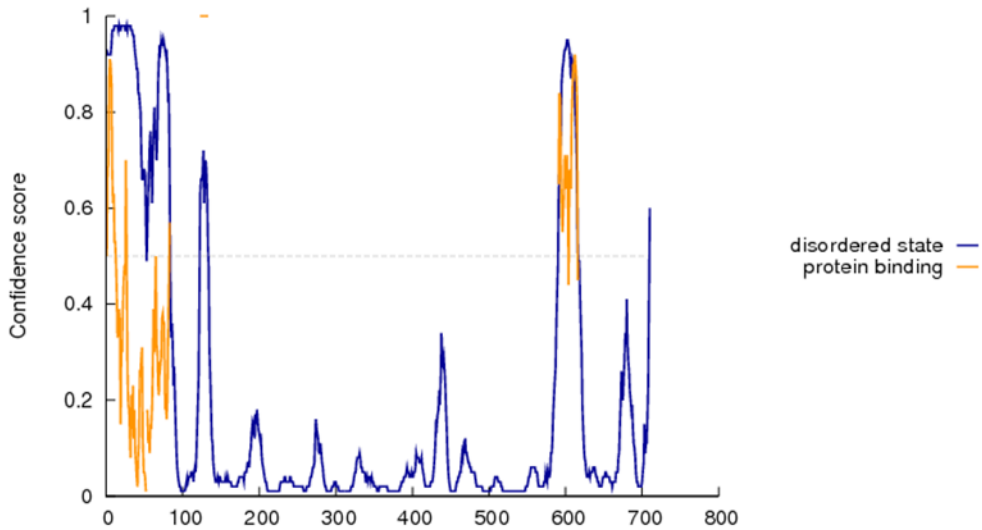

Figure 1. Intrinsic disordered profile of buffalo Mx2 protein predicted by DISOPRED3 program. Amino acids sequence is in x-axis and y-axis represents the level of confidence with grey line as threshold value for disordered region.

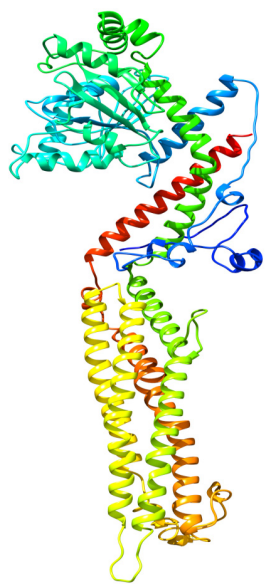

Figure 2a. 3D structure of Model 4 generated by I-TASSER.

Figure 2b. 3D structure of refined Model generated in Modeller 9.19 (both the models are viewed and presented from UCSF Chimera).

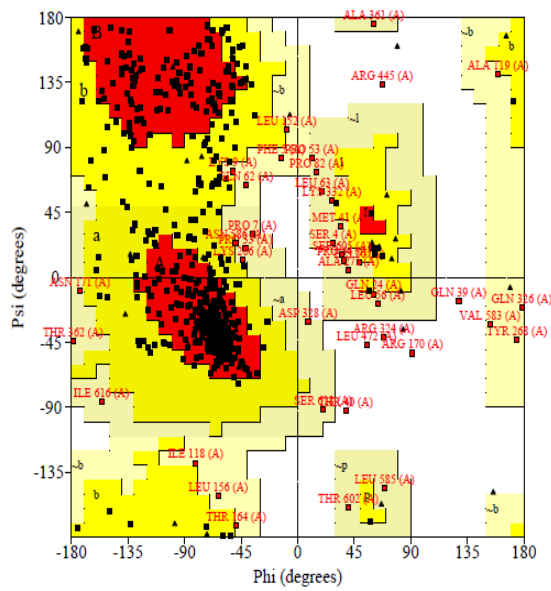

Figure 2c. Ramachandran plot of Model 4 of I-TASSER.

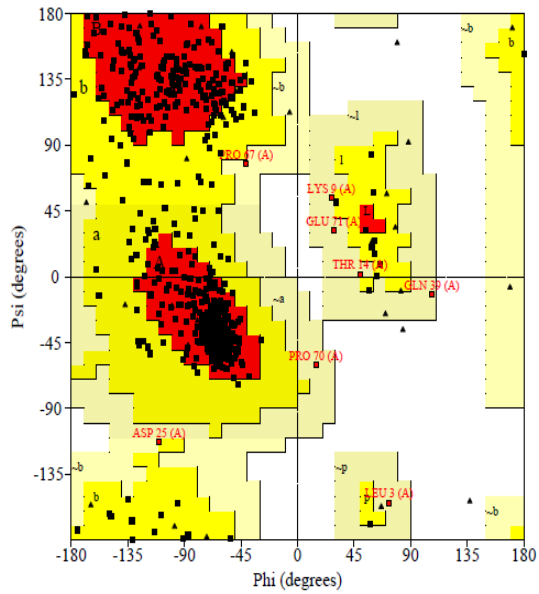

Figure 2d. Ramachandran plot of refined Model (Generated by PROCHECK program).

Figure 3. Superimposed 3D structure of refined buffalo Mx2 protein model (red colour) and human MxB protein model (Chain A of PDB ID: 4WHJ) (blue colour).

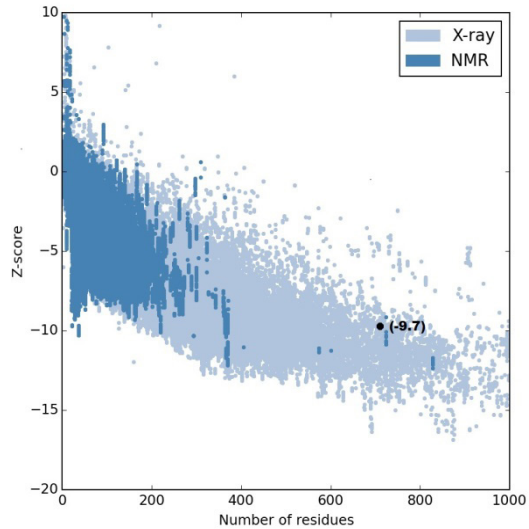

Figure 4a. Refined model of buffalo Mx2 protein shown as black spot with Z-score -9.7.

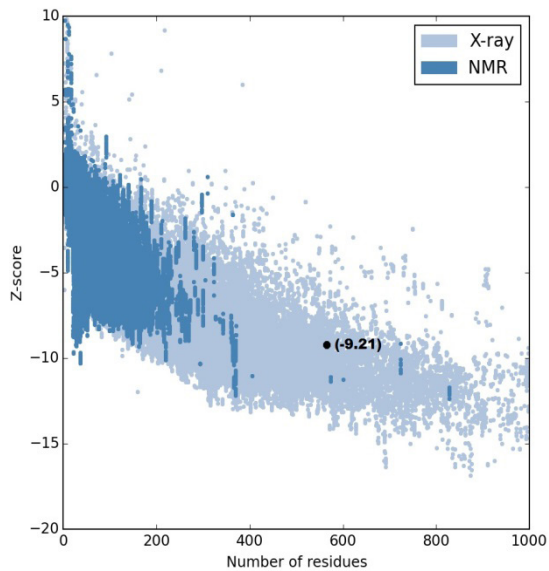

Figure 4b. PDB model of human MxB protein (4WHJ) represented as black spot with Z-score -9.21.

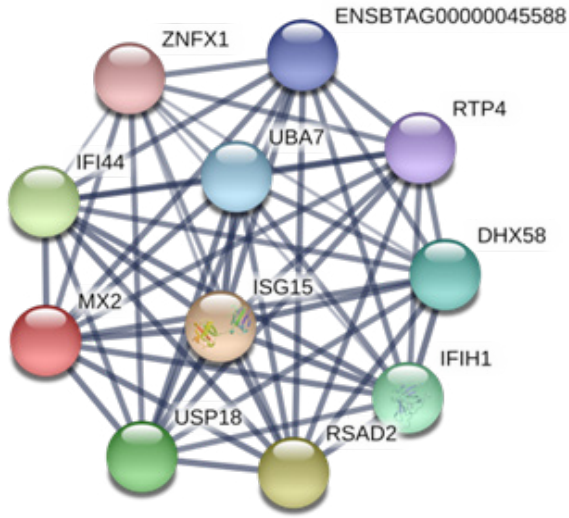

Figure 5. Interactions of buffalo Mx2 protein with other ten proteins predicted by online tool STRING 10.5. Solid lines joining two proteins indicate a reliable interaction.

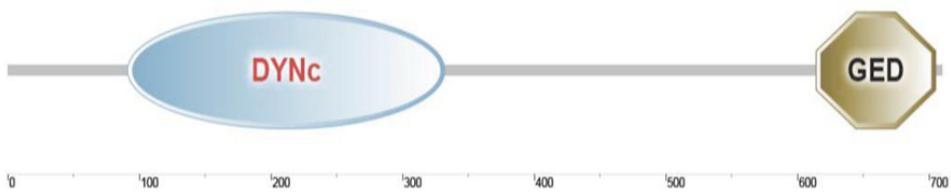

Figure 6. Domains of buffalo Mx2 protein identified by using SMART. DYNc region represents the Dynamin, GTPase domain and GED region represents the Dynamin GTPase effector domain.

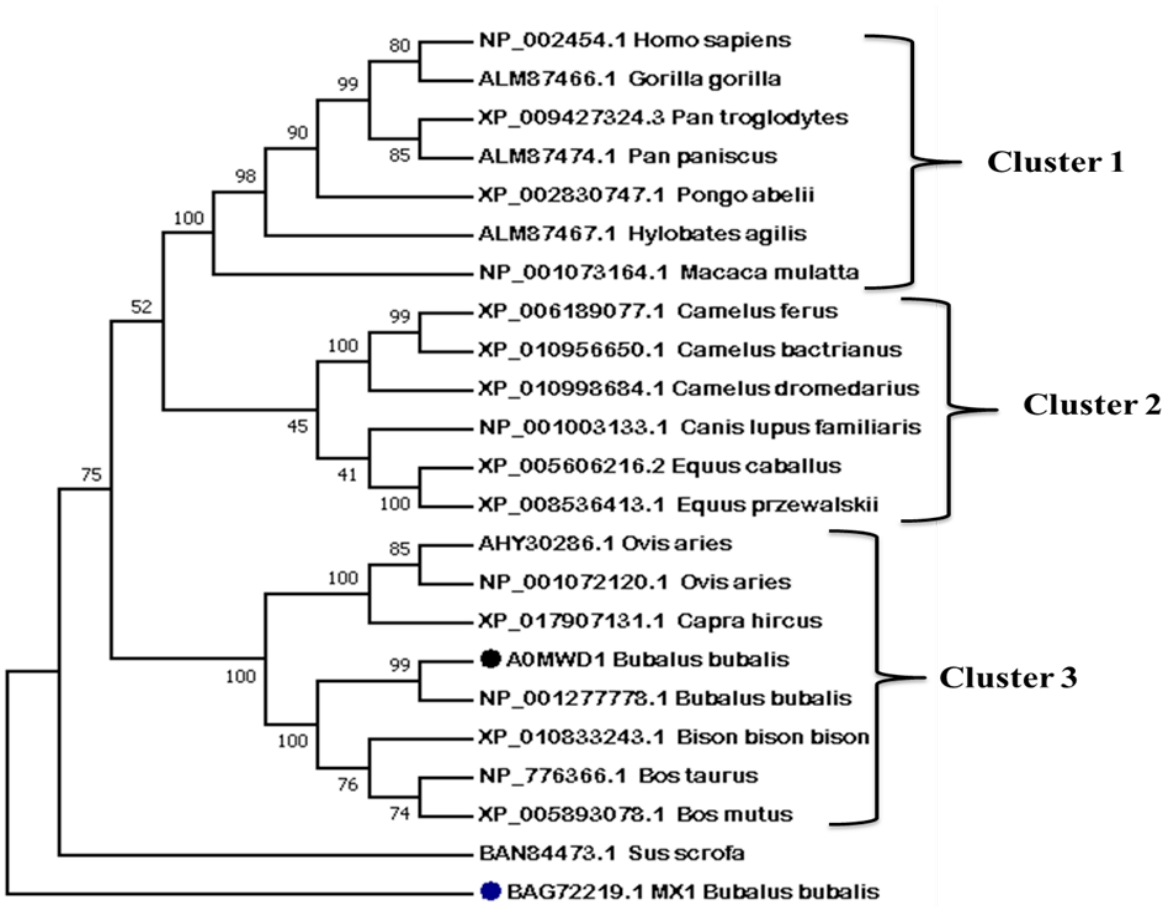

Figure 7. Phylogenetic tree of Mx2 protein of buffalo and other mammalian species. The tree was constructed in MEGA7 software by UPGMA method. Different mammalian species along with GenPept IDs are shown. Numbers along the branches refer to the bootstrap value (percentage of confidence). Buffalo Mx2 protein sequence (A0MWD1) used in this study is highlighted with black solid circle and outgrouped sequence is represented with blue solid circle.

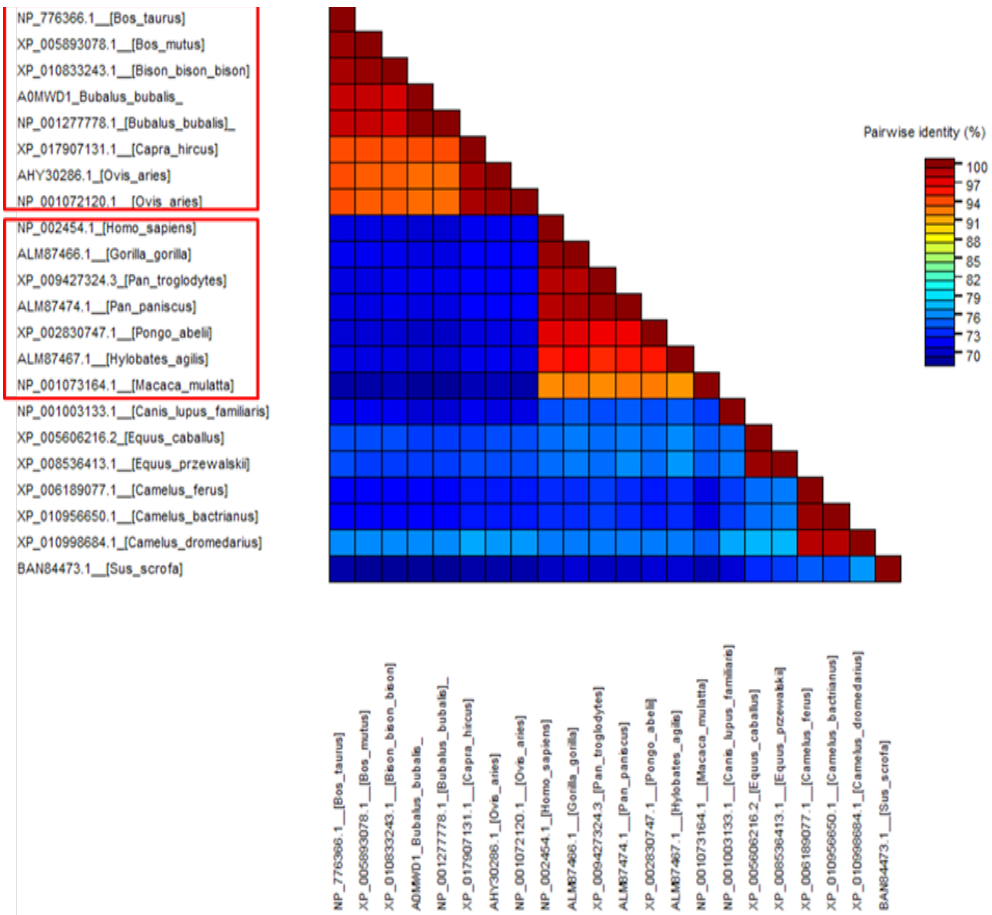

Figure 8. Pairwise identity matrix of Mx2 protein sequence of various mammalian species. A colour-coded pairwise identity matrix was generated from diverse mammalian Mx2 protein sequences (species name and GenPept IDs are shown) using SDT program. Each coloured cell represents a percentage of identity score between two sequences (one indicated horizontally the other vertically). Species forming different subgrouped are demarcated by red box.

Protein interactions and structural organization

The probable protein-protein interactions in which Mx2 protein of buffalo might participate was analysed using STRING 10.5 tool. This online tool does not contain buffalo in its search options for species selection; therefore, analysis was done by selecting the closest organism *Bos taurus*. The top 10 proteins that interacted with Mx2 were ISG15, RSAD2, IFI44, USP18, IFIH1, DHX58, UBA7, ENSBTAG00000045588, RTP4 and ZNFX1 (Figure 5). All the 10 predicted interactions had good score ranging from 0.938 to 0.828, indicating a reliable interaction among them (Table 2). The proteins ISG15 and RSAD2 were known to have antiviral activity like Mx2, the proteins IFI44, IFIH1and ENSBTAG00000045588 were uncharacterized and the rest were having different functions (Table 2). The *in silico* interactions suggested that buffalo Mx2 protein not only participate in host viral interaction but also in many other essential cellular events.

The SMART-Simple Modular Architecture Research Tool confidently predicted two domains in Mx2 protein of buffalo. The first domain was Dynamin, GTPase (DYNc) which was located from residue 91 to 332 and the other is Dynamin GTPase effector domain (GED) that starts at position 614 and ends at 705 (Figure 6). The presence of these two domains indicated that buffalo Mx2 protein belongs to dynamin superfamily like other Mx proteins and apparently involves in endocytosis besides its antiviral activity, since most of the proteins belonging to dynamin family participate in various processes of endocytosis (Praefcke and Mahon, 2004; Ferguson, *et al*., 2012).

Phylogenetic analysis

The amino acid substitution model Jones-Taylor-Thornton (Jones *et al*., 1992) with gamma distribution (JTT+G) was estimated to be the best fit model in MEGA 7 software on the basis of BIC. The phylogenetic tree constructed by UPGMA (Sneath and Sokal, 1973) method revealed that Mx2 protein of buffalo has the same ancestor with that of cattle, sheep and goat because a distinguished clade (cluster 3) was formed by the members of bovidae family (Figure 7). Typically, the mammal from hominidae and hylobatidae family formed one clade (cluster 1) but it was outlandish that members from other families such as equidae, camelidae and canidae formed one clade (cluster 2) except suidae which formed an individual clade (Figure 7). The phylogenetic tree also revealed that the ancestry of buffalo Mx2 protein is closer to ruminants than non -ruminants and far away from human and other primates. The pairwise identity of Mx2 protein sequence estimated in SDT suggested that heterogonous mammalian Mx2 proteins are of the same molecular group as they exhibited >69% identity among themselves (Figure 8). Moreover, based on sequence identity, the Mx2 protein from bovidae family could be clustered as one subgroup while hominidae and hylobatidae were clustered as another subgroup as they shared >91% pairwise identity among the sequences (Figure 8). Other sequences were ungrouped and analyses with more number of sequences from related species are needed for further subgrouping them. This finding additionally supports the phylogenetic tree which is suggestive of molecular relatedness of Mx2 proteins belonging to diverse mammalian species.

Overall, the *in-silico* experimental results manifested many hidden functional, structural and evolutionary characteristics of Mx2 protein of *Bubalus bubalis*. Although the characterization of the studied protein is theoretical in nature but all the information produced by computational

approach could assist future *in-vitro* and *in-vivo* experiments and further characterization of buffalo Mx2 protein.

ACKNOWLEDGMENTS

The authors are thankful to the Dean, College of Veterinary Science, Guwahati as well as acknowledge the Bioinformatics Infrastructure Facility, C.V.Sc., Khanapara, Guwahati, sponsored by Department of Biotechnology, Ministry of Science and Technology, Govt. of India for providing facilities to conduct the research.

REFERENCES

Aebi, M., J. Fah, N. Hurt, C.E. Samuel, D. Thomis, L. Bazzigher, J. Pavlovic, O. Haller and P. Staeheli. 1989. cDNA structures and regulation of two interferon-induced human Mx proteins. *Mol. Cell. Biol*., **9**: 5062-5072.

Babiker, H.A.E., T. Saito, Y. Nakatsu, S. Takasuga, M. Morita, Y. Sugimoto, J. Ueda and T. Watanabe. 2016. Molecular cloning, polymorphism, and functional activity of the bovine and water buffalo Mx2 gene promoter region. *SpringerPlus*, **5**: 2109.

Babiker, H.A., Y. Nakatsu, K. Yamada, A. Yoneda, A. Takada, J. Ueda, H. Hata and T. Watanabe. 2007. Bovine and water buffalo Mx2 genes: Polymorphism and antiviral activity. *Immunogenetics*, **59**: 59-67.

Benkert, P., S.C.E. Tosatto and D. Schomburg. 2008. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Structure, Function, and Bioinformatics*, **71**: 261-277.

Blom, N., S. Gammeltoft and S. Brunak. 1999. Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol*., **294**: 1351-1362.

Cour, T. la, L. Kiemer, A. Molgaard, R. Gupta, K. Skriver and S. Brunak. 2004. Analysis and prediction of leucine-rich nuclear export signals. *Protein Engineering Design and Selection*, **17**(6): 527-536.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*., **32**: 1792-1797.

Ellinwood, N.M., J.M. McCue, P.W. Gordy and R.A. Bowen. 1998. Cloning and characterization of cDNAs for a bovine (*Bos taurus*) Mx protein. *J. Interf. Cytok. Res*., **18**: 745-755.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**: 783-791.

Ferguson, S.M. and P. De Camilli. 2012. Dynamin, a membrane remodelling GTPase. Nature Reviews. *Mol. Cell. Biol*., **13**: 75-88.

Geourjon, C. and G. Deleage. 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci*., **11**: 681-684.

Gill, S.C. and P.H. von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem*., **182**: 319-326.

Guruprasad, K., B.V. Reddy and M.W. Pandit. 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. *Protein Engineering Design and Selection*, **4**: 155-161.

Hicks, B.A., S.J. Etter, K.G. Carnahan, M.M. Joyce, A.A. Assiri, S.J. Carling, K. Kodali, G.A. Johnson, T.R. Hansen, M.A. Mirando, G.L. Woods, D.K. Vanderwall and T.L. Ott. 2003. Expression of the uterine Mx protein in cyclic and pregnant cows, gilts, and mares. *J. Anim. Sci*., **81**: 1552-1561.

Horisberger, M.A. and M.C. Gunst. 1991. Interferon-induced proteins: Identification of Mx proteins in various mammalian species. *Virology*, **180**: 185-190.

Horton, P., K.J. Park, T. Obayashi, N. Fujita, H. Harada, C.J. Adams-Collier and K. Nakai. 2007. WoLF PSORT: protein localization predictor. *Nucleic Acids Res*., **35**: W585-W587.

Ikai, A.J. 1980. Thermostability and aliphatic index of globular proteins. *J. Biochem*., **88**: 1895-1898.

Johansen, M. Bo, L. Kiemer and S. Brunak. 2006. Analysis and prediction of mammalian protein glycation. *Glycobiology*, **16**: 844-853.

Jones, D.T. and D. Cozzetto. 2015. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*, **31**: 857-863.

Jones, D.T., W.R. Taylor and J.M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Bioscie.*, **8**: 275-282.

Julenius, K. 2007. NetCGlyc 1.0: Prediction of mammalian C-mannosylation sites. *Glycobiology*, **17**: 868-876.

Kiemer, L., J.D. Bendtsen and N. Blom. 2005. NetAcet: Prediction of N-terminal acetylation sites. *Bioinformatics*, **21**: 1269-1270.

Kojima, T., K. Oshima, H. Watanabe and M. Komatsu. 2003. The bovine Mx1 gene: Characterization of the gene structure, alternative splicing, and promoter region. *Biochem. Genet*., **41**: 375-390.

Krogh, A., B. Larsson, G. von Heijne and E.L.L. Sonnhammer. 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol*., **305**: 567-580.

Kumar, S., G. Stecher and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*., **33**: 1870-1874.

Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol*., **157**: 105-132.

Laskowski, R.A., J.A. Rullmannn, M.W. MacArthur, R. Kaptein and J.M. Thornton. 1996. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**: 477-486.

Laskowski, R.A., M.W. MacArthur, D.S. Moss and J.M. Thornton. 1993. PROCHECK - A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*., **26**: 283-291.

Letunic, I. and P. Bork. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*., **46**: D493-D496.

Lindenmann, J., C.A. Lane and D. Hobson. 1963. The resistance of A2G mice to myxoviruses. *J. Immunol*., **90**: 942-951.

Matsuyama, S., T. Kojima, S. Kato and K. Kimura. 2012. Relationship between quantity of IFNT estimated by IFN-stimulated gene expression in peripheral blood mononuclear cells and bovine embryonic mortality after

AI or ET. *Reprod. Biol. Endocrinol*., **10**: 21.

Melen, K., T. Ronni, B. Broni, R.M. Krug, C.H. Von Bonsdorff and I. Julkunen. 1992. Interferon induced Mx proteins form oligomers and contain a putative leucine zipper. *J. Biol. Chem*., **267**: 25898-25907.

Muhire, B.M., A. Varsani and D.P. Martin. 2014. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PloSOne*, **9**: e108277.

Muller, M., E.L. Winnacker and G. Brem. 1992. Molecular cloning of porcine Mx cDNAs: New members of a family of interferon-inducible proteins with homology to GTP-binding proteins. *J. Interferon Res*., **12**: 119-129.

Nakatsu, Y., K. Yamada, J. Ueda, A. Onogi, G.P. Ables, M. Nishibori, H. Hata, A. Takada, K. Sawai, Y. Tanabe, M. Morita, M. Daikohara and T. Watanabe. 2004. Genetic polymorphisms and antiviral activity in the bovine MX1 gene. *Anim. Genet*., **35**: 182-187.

Ott, T.L., J. Yin, A.A. Wiley, H.T. Kim, B. Gerami-Naini, T.E. Spencer, F.F. Bartol, R.C. Burghardt and F.W. Bazer. 1998. Effects of the estrous cycle and early pregnancy on uterine expression of Mx protein in sheep (*Ovis aries*). *Biol. Reprod*., **59**: 784-794.

Pavlovic, J., T. Zurcher, O. Haller and P. Staeheli. 1990. Resistance to influenza virus and vesicular stomatitis virus conferred by expression of human MxA protein. *J. Virol*., **64**: 3370-3375.

Pavlovic, J., A. Schroder, A. Blank, F. Pitossi and P. Staeheli. 1993. Mx proteins: GTPases involved in the interferon-induced antiviral state. *Ciba Foundation Symposium*, **176**: 233-243.

Petersen, T.N., S. Brunak, G. von Heijne and H. Nielsen. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods*, **8**: 785-786.

Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin. 2004. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem*., **25**: 1605-1612.

Pierleoni, A., P.L. Martelli, P. Fariselli and R. Casadio. 2006. BaCelLo: A balanced subcellular localization predictor. *Bioinformatics*, **22**: e408-e416.

Praefcke, G.J. and H.T. McMahon. 2004. The dynamin superfamily: Universal membrane tubulation and fission molecules. *Nat. Rev. Mol. Cell Biol*., **5**: 133-147.

Roy, A., J. Yang and Y. Zhang. 2012. Cofactor: An accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res*., **40**: W471-W477.

Sali, A. and T.L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol*., **234**: 779-815.

Sippl, M.J. 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins*, **17**: 355-362.

Sneath, P.H.A. and R.R. Sokal. 1973. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. W.H. Freeman, San Francisco, USA.

Spencer, T.E., O. Sandra and E. Wolf. 2008. Genes involved in conceptus - Endometrial interactions in ruminants: Insights from reductionism and thoughts on holistic approaches. *Reproduction*, **135**: 165-179.

Staeheli, P. and J.G. Sutcliffe. 1988. Identification of a second interferon-regulated murine Mx

gene. *Mol. Cell Biol*., **8**: 4524-4528.

Staeheli, P., F. Pitossi and J. Pavlovic. 1993. Mx proteins: GTPases with antiviral activity. *Trends Cell Biol*., **3**: 268-272.

Staeheli, P., O. Haller, W. Boll, J. Lindenmann, C. Weissmann. 1986. Mx protein: Constitutive expression in 3T3 cells transformed with cloned Mx cDNA confers selective resistance to influenza virus. *Cell*, **44**: 147-158.

Steentoft, C., S.Y. Vakhrushev, H.J. Joshi, Y. Kong, M.B. Vester-Christensen, K.T. Schjoldager, K. Lavrsen, S. Dabelsteen, N.B. Pedersen, L. Marcos-Silva, R. Gupta, E.P. Bennett, U. Mandel, S. Brunak, H.H. Wandall, S.B. Levery and H. Clausen. 2013. Precision mapping of the human O-GalNAc glycoproteome through simple cell technology. *EMBO J*., **32**: 1478-1488.

Szklarczyk, D., J.H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N.T. Doncheva, A. Roth, P. Bork, L.J. Jensen and C. von Mering. 2017. The STRING database in 2017: Quality - controlled protein - protein association networks, made broadly accessible. *Nucleic Acids Res*., **45**: D362-D368

Wiederstein, M. and M.J. Sippl. 2007. ProSA-web: interactive web service for the recognition of errors in three - Dimensional structures of proteins. *Nucleic Acids Res*., **35**: W407-W410.

Wu, S. and Y. Zhang. 2007. LOMETS: A local meta - Threading-server for protein structure prediction. *Nucleic Acids Res*., **35**: 3375-3382.

Yamada, K., Y. Nakatsu, A. Onogi, J. Ueda and T. Watanabe. 2009. Specific intracellular localization and antiviral property of genetic and splicing variants in bovine Mx1. *Viral Immunol*., **22**: 389-395.

Yang, J. and Y. Zhang. 2015. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res*., **43**: W174-W181.

Zhang, Y. and J. Skolnick. 2004. SPICKER: Approach to clustering protein structures for near-native model selection. *J. Comput. Chem*., **25**: 865-871.

Zhang, Y. 2009. I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins*, **77**: 100-113.

Zhang, Y. and J. Skolnick. 2005. TM-align: A protein structure alignment algorithm based on TM-score. *Nucleic Acids Res*., **33**: 2302-2309.

Zurcher, T., J. Pavlovic and P. Staeheli. 1992. Mouse Mx2 protein inhibits vesicular stomatitis virus but not influenza virus. *Virology*, **187**: 796-800.