



In silico* characterization of interferon-stimulated gene (ISG15) as a biomarker for early pregnancy diagnosis in *Bubalus bubalis

KANISHT BATRA¹ and TRILOK NANDA²

Lala Lajpat Rai University of Veterinary and Animal Sciences, Hisar, Haryana 125 004 India

Received: 8 August 2018; Accepted: 6 September 2018

ABSTRACT

Reproductive management of livestock is the foremost requirement for increasing production from any dairy animal. High reproductive efficiency can be achieved by implementation of one of the crucial management practices which involves early differentiation of pregnant and non-pregnant animals. The early and precise pregnancy diagnosis is the major problem in *Bubalus bubalis* (buffalo) due to absence of estrous signs instead of normal cyclicity of reproductive organs. Therefore, in the present study, one of the conceptus derived protein interferon stimulated protein (15 kDa) released during 18 to 21 days after insemination in response to implantation of embryo was *in silico* characterized. The protein expressed by interferon-stimulated gene-15 (ISG15) was analyzed by isolation and sequencing of coding region of mRNA. The primary and secondary structures were predicted from the protein sequence to decipher the interaction of ISG15 protein with other molecules. The functional characterization identifies various motifs present in ISG15 protein which are responsible for its interaction with other proteins. Physiochemical properties predicted the ISG15 protein nature during *in vitro* conditions which are required for any of the assays development. In addition, immunogenic studies revealed ISG15 protein is strongly antigenic in nature and can be used for antibody production. In conclusion, ISG15 protein expression from buffalo is a good indicator of conceptus implantation and has suitable properties for being used as target to develop early pregnancy diagnostic kits.

Key words: Biomarker, *Bubalus bubalis*, Characterization, *In silico*, ISG15, Pregnancy detection

Interferon stimulated protein is one of the most expressed proteins which have been reported not only in antiviral host response but also during early pregnancy within the uterus of bovine (Austin *et al.* 1996, Hansen *et al.* 1997), ovine (Johnson *et al.* 1999), porcine (Joyce *et al.* 2002), murine (Austin *et al.* 2003), human and baboon (Bebington *et al.* 1999) species. Among these, Buffalo constitutes one of the most important dairy animals present in Indian subcontinent and has a large contribution to world total milk production of about 30 million tonnes annually (<http://www.dahd.nic.in>). In spite of availability of many advanced technological tools, it is experiencing problems related to reproduction such as high calving interval, silent heat, late puberty, high incidence of anestrus and lack of early pregnancy diagnosis tools. There are several methods of pregnancy diagnosis which are widely adopted such as direct palpation methods, ultrasound, progesterone based assay but current requirement is to identify a molecule which can replace these low sensitive and less accurate conservative methods. Several attempts have been made to identify the protein molecule which can be used for early

detection of pregnancy (Singh *et al.* 2005, Balhara *et al.* 2014). Therefore, a proper protein expressed by interferon-stimulated gene (ISG15) which is specifically up or downregulated in response to conceptus implantation can act as suitable candidate for pregnancy detection.

ISG15 protein is released in response to Interferon tau (IFN- τ) which is secreted by the conceptus on day 18 to 21 of pregnancy (Johnson *et al.* 1998, Perry *et al.* 1999, Thatcher *et al.* 2001). However, concentration of IFN- τ in blood is extremely low whereas ISG15 protein level is at measurable level. The ISG15 protein is responsible for modification of many proteins by attaching covalently to them such as ubiquitin and responsible for regulating their activity similar to protein phosphorylation or acetylation (Loeb and Haas 1992, Hamerman *et al.* 2002). The ISG15 protein can degrade cytosolic uterine proteins (e.g. receptors, enzymes and transcription factors of regulating genes) that are detrimental for fetal/embryo survival (Johnson *et al.* 1998). Also, ISG15 protein plays a key role at the embryo-maternal interface (Austin *et al.* 2003, Joyce *et al.* 2005). The expression of the ISG15 gene in response to the embryo plays an important biological role in uterine receptivity as well as conceptus implantation (Bazer *et al.* 2009, Hansen *et al.* 1999, Spencer *et al.* 2008). A transcriptome analysis in the cattle endometrium has also

Present address: ¹Research Associate (drkanishtbatra@gmail.com), ICAR-National Research Center on Equines, Hisar, Haryana. ²Professor (nandatriloc@rediffmail.com), Department of Animal Biotechnology.

identified ISG15 gene as one of the differentially expressed genes in response to conceptus derived IFN τ (Forde *et al.* 2011). Recently, a significant higher expression of ISG15 mRNA was found on day 16 ($P < .05$) and day 18 ($P < .05$) of pregnancy in nulliparous heifers (Soumya *et al.* 2016). Till date, no tool was available which can employ this molecule as a suitable marker for early pregnancy detection. The present study was planned to find out the suitability of ISG15 protein for development of early pregnancy diagnostic kit. Therefore, in the present study, ISG15 gene was isolated from *Bubalus Bubalis* and reverse genetics approach was applied to decipher its characteristics by *in silico* approach.

MATERIALS AND METHODS

Collection of samples: All procedures were conducted in accordance with guidelines approved by the Institutional Animal Ethics Committee (IAEC), registered as 1669/GO/ReBiBt/S/12/CPCSEA dated 6.12.2012. meeting dated 6th February 2017. Blood samples were collected from Murrah buffaloes from organized farm. Animal showing estrus signs were inseminated and used for collection of blood sample which was marked as day 0. Samples collected at 0th, 18th and 21st day after inseminations were immediately processed for RNA extraction.

RNA isolation: Extraction of total RNA from the blood samples was carried out using TRIzol Reagent (Ambion). Briefly, 600 μ l of blood was mixed with 800 μ l of TRIzol reagent and vortexed uniformly. Chloroform was added for separation into organic phase and aqueous phase. The upper aqueous phase aqueous was precipitated by adding 500 μ l isopropanol and incubation overnight at -20°C . On next day, the RNA was pelleted at 12,000 rpm for 20 min and washed with 1 ml of 70% ethanol. The pellet was resuspended in 20 μ l of nuclease free water. Total RNA was reverse transcribed using Revertaid cDNA Synthesis kit (Thermo) as per manufacturer's instructions.

Primer designing: Primers were designed using Primer BLAST software from NCBI using the gene sequence of accession no. NM_001291322.1 (buffalo). Primers (F-5' CACCATGGGCGGGGACCTGAAGG and R-3' CTACCCACCCCGCAGACGTAGAT) (Batra *et al.* 2018a) were designed with CACC overhang at 5' end for cloning in championTM pET100 directional TOPO[®] expression kit as per manufacturer's instructions (Invitrogen).

Conventional PCR: PCR was performed in Thermal cycler (Veriti Applied Biosystem) in 25 μ l reaction containing 5 μ l of template cDNA, 0.4 μ M of forward and reverse primer, using high fidelity Fusion Taq mastermix (2x concentration) with HF buffer (NEB). Cyclic conditions for PCR were optimized as initial denaturation at 94°C for 3 min and 40 cycles of denaturation at 94°C for 30 sec, annealing at 55°C for 45 sec, elongation at 72°C for 2 min followed by final elongation at 72°C for 10 min (Batra *et al.* 2017).

Cloning of ISG15: Amplified ISG15 product was mixed with pET100 vector for topoisomerase ligation reaction.

Ligation mixture was transformed inside one shot[®] TOP 10 chemically competent *E. coli* (Invitrogen). Selection of recombinant clones was done on basis of ampicillin resistant gene. Plasmids were isolated from positive clones using plasmid isolation kit (Invitrogen) as per manufacturer's instructions and confirmed by gene specific and vector specific primers.

Sequencing of plasmid DNA: Recombinant plasmids were sequenced using Big Dye terminator v3.1 cycle sequencing kit (Applied Biosystems) on an ABI 3130XL genetic analyser. The contig sequence obtained was aligned with the available sequences in the GenBank using MEGA 6.0 programme and used to construct phylogenetic tree to show relatedness between all available ISG15 sequences in different species. This sequence was translated into amino acid sequences using expasy translate online tool (<http://web.expasy.org/translate>) and then further characterized.

Prediction of primary structure: The protein primary structure of ISG15 protein was predicted using academic version of MODELER 9.8 (<http://www.salilab.org/modeller>) (Sali and Blundell 1993). The probable density function (pdf) value of ISG15 protein was estimated by Modeller 9.8, which signify the spatial restraint used in restraining the model structure (Lovell *et al.* 2002). A 3D model of ISG15 protein was obtained by optimization of the molecular pdf value with the violation of restraints as little as possible. This model generated was visualized in swiss pdb viewer (Guex and Peitsch 1997). Sequence alignment between buffalo protein and other species proteins was done using T coffee program (Notredame *et al.* 2011). Model was also generated using automated Swiss prot modeler (Arnold *et al.* 2006) which listed the global model quality estimate and the QMEAN score of different models. The structure was confirmed using Procheck and verifies with VERIFY3D (Luthy *et al.* 1992). RAMPAGE was used to check the stereochemical properties of ISG15 protein structure (Ramachandran *et al.* 1963).

Prediction of secondary structure: The secondary structure of ISG15 protein was predicted using SOPMA (Self-Optimized Prediction Method with Alignment) tool (Sapay *et al.* 2006). The graphical representation of secondary structure was predicted using Protein Homology Recognition Engine v2.0 (PHYRE) (Kelley *et al.* 2015) which predicted the presence of alpha-helices, beta-strands and disordered regions in ISG15 protein in graphical form with a color-coded confidence bar.

Prediction of linear motif and functional characterization: The Eukaryotic linear Motiff (ELM) (Dinkel *et al.* 2016) prediction tool was used for prediction of motiffs in ISG15 protein sequence by scanning the sequences with the regular expressions defined in ELM.

Physicochemical properties of protein: The physicochemical properties were calculated using Expasy's Prot-Param (Gasteiger *et al.* 2005) (<http://us.expasy.org/tools/protparam.html>) prediction server. This predicted different properties of ISG15 protein such as molecular weight, theoretical isoelectric point (pI), total number of

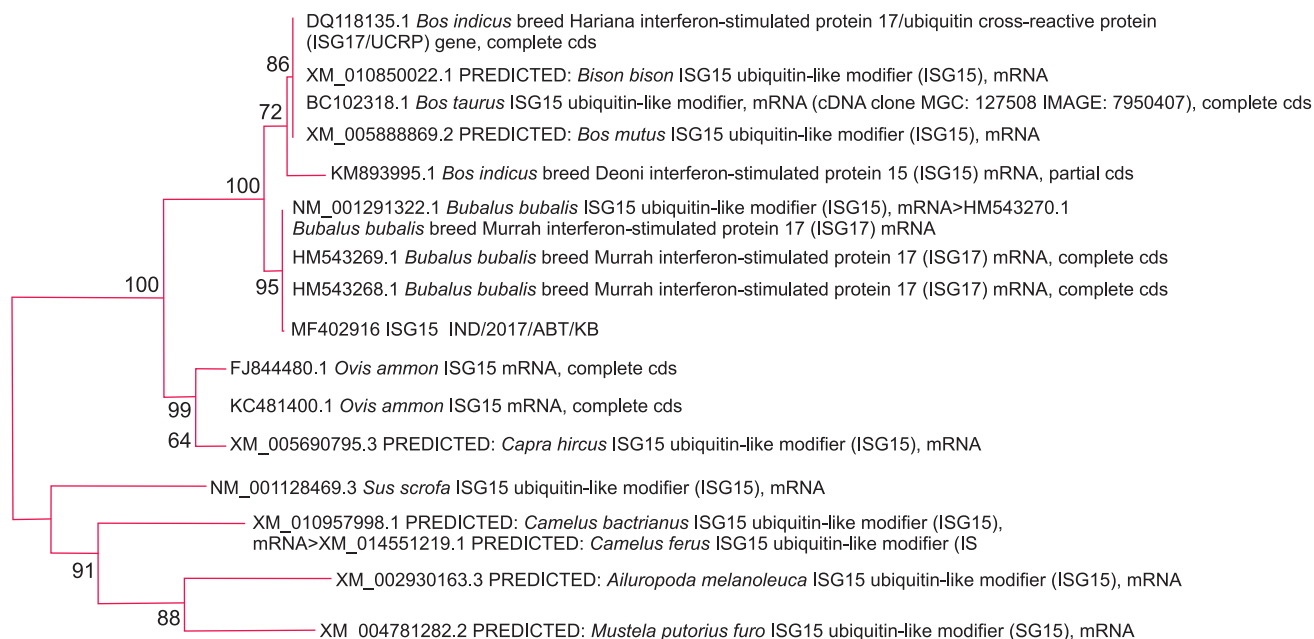


Fig. 1. Unrooted neighbor joining phylogenetic tree drawn from complete ORF sequences of ISG15

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).

Prediction of accessible surface area (ASA): The accessible surface area of the ISG15 protein was predicted through NetSurfP server (Petersen *et al.* 2009) of ExPaSy software.

Predicting phosphorylation sites and glycosylation sites: NetPhos server tool was used to predict different phosphorylation sites present in ISG15 protein capable of binding phosphate group (Blom *et al.* 1999). The glycosylation sites were predicted using GLCO ep software (<http://www.imtech.res.in/raghava/glycoep/submit.html>) using ISG15 sequence (Chauhan *et al.* 2013).

Prediction of disulfide bridges: CYS_REC tool (http://www.softberry.com/cgi-bin/programs/propt/cys_rec) was utilized for prediction of disulfide bridges which recognized the positions of cysteines present and computes the most probable S-S bond pattern of the pairs in ISG15 protein sequence.

Prediction of methylation and acetylation: Prediction of potential methylation and acetylation of ISG15 protein sequence was done using *in silico* tool PLMLA (Prediction of potential lysine methylation and lysine acetylation) (Shi *et al.* 2012).

Conserved domain identification for function prediction: The ISG15 protein sequences were used as input using NCBI CDD-BLAST tool (available at <http://www.ncbi.nlm.nih.gov>) for searching conserved domains.

Prediction of immunogenic sites: B-cell epitope databases were utilized in order to evaluate the existing epitopes on ISG15 protein. The antigenicity of protein due to linear as well conformational epitopes were predicted

using IEDB analysis software (Larsen *et al.* 2006) by providing the ISG15 sequence.

RESULTS AND DISCUSSION

IFN τ release from a blastocyst (day 15 to 16 after breeding) clearly alters endometrial gene expression, which allows the release of ISG15. This molecule has revealed significant higher expression ($P < 0.05$) in pregnant buffaloes on 18 to 21 days post AI than non pregnant group (Han *et al.* 2006, Meyerholz *et al.* 2016). Therefore, ISG15 protein with remarkable property of expression over a considerable period along with no residue after pregnancy was taken into consideration to find its suitability for early pregnancy diagnosis biomarker.

Recently, expression and molecular cloning of Interferon stimulated genes in buffalo (*Bubalus bubalis*) had been reported. They demonstrated the expression dynamics of ISG15, OAS1, and Mx2 gene and revealed their potency towards ideal biomarker for detection of early pregnancy in buffalo (Thakur *et al.* 2017, Batra *et al.* 2018b, c). Soumya *et al.* (2016) had also observed the interferon-stimulated gene 15 (ISG15) as one of the major gene stimulated by interferon tau, the maternal recognition of pregnancy signal in ruminants. In the present study, on the basis of semi quantitative RT-PCR, specific to ISG15 mRNA, it was found that there was specific amplification of 466 bp product during 18 to 21 day of gestation which was indicative of increased expression level of ISG15 gene. These results of present study obtained were parallel to findings by different authors (Green *et al.* 2010, Matsuyama *et al.* 2012, Kizaki *et al.* 2013) that expression level of ISG15 gene was higher during 18 to 21 days of pregnancy in blood of animals.

ISG15 protein is a novel molecule which needs to be

deciphering more for finding its suitability to use a marker molecule in early pregnancy detection. In the present study, complete ORF region of ISG15 gene was amplified and cloned in pET 100 vector.

The plasmid was sequenced and contig sequence obtained was submitted in gene bank with accession number MF402916 which was further aligned with the available sequences in the GenBank using MEGA 6.0 programme. Phylogenetic analysis revealed the sequence derived from cloned ORF region has considerable nucleotide differences and have 100% similarity with buffalo sequence, 98% *Bos taurus*, 94% *Ovis aries* and 93% *Capra hircus* (Fig. 1). The amino acid sequence derived was 98% identical with 98% *Bos taurus*, 89% *Capra hircus* and 81% *Ovis aries*. The present study of ISG15 gene revealed little structural differences which indicate the suitability of molecule for

acting as biomarker in other species also (Joyce *et al.* 2005).

ISG15 protein becomes conjugated to Ubiquitin-activating enzyme 1-like protein (UBE1L) in response to release of IFNs (Rempel *et al.* 2005). UBE1L mRNA was localized within the luminal and glandular epithelium and more diffusely within the stroma similar to ISG15 mRNA in endometrial sections from pregnant cows. This provides the insight that not alone ISG15 protein but proteins conjugated to this are also stimulated by release of INF- τ during conceptus elongation (Malakhov *et al.* 2003, Jain *et al.* 2012). Therefore, primary and secondary structure prediction using transcriptome study was done for prediction of interaction with other molecules during conceptus implantation. The modeled primary structure deduced by structural prediction algorithms is required to evaluate the conformational differences with other species, its cross

Table 1. Different ELM motifs present in ISG15 sequences

Elm Name	Instances (Matched sequence)	Position	Elm description	Cell compartment
CLV_PCSK_SKI1_1	REVLQ	7–11	Subtilisin/kexin isozyme-1 (SKI1) cleavage site ([RK]-X-[hydrophobic]-[LTKF]-I-X)	Endoplasmic reticulum lumen, Golgi apparatus, extracellular
DEG_Nend_Nbox_1	LA	1–2	N-terminal motif that initiates protein degradation by binding to the N-box of N-recognins	cytosol
DOC_MAPK_MEF2A_6	RAGSTVLLVV	22–31	A kinase docking motif that mediates interaction towards the ERK1/2 and p38 subfamilies of MAP kinases	Cytosol, Transcription factor complex, Nucleus
DOC_WW_Pin1_4	GRSSPY	44–49	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase.	Cytosol, Nucleus
LIG_14-3-3_CanoR_1	RAGSTVL	22–28	Canonical Arg-containing phospho-motif mediating a strong interaction with 14-3-3 proteins.	Cytosol, Internal side of plasma membrane, Nucleus
LIG_FHA_1	GSTVLLV	24–30	Phosphothreonine motif binding site	Nucleus
LIG_FHA_2	KQTVAEL	54–60	Phosphothreonine motif binding site	Nucleus, Replication fork
LIG_LIR_Gen_1	TVFMNL	100–105	Canonical LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy	Cytosol, Cytoplasmic side of late endosome membrane
LIG_LIR_LC3C_4	TVLLVV	26–31	Non-canonical variant of the LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy	Cytosol, Cytoplasmic side of late endosome membrane
LIG_SH2_STAT3	YEVQ	49–52	YXXQ motif found in the cytoplasmic region of cytokine receptors that bind STAT3 SH2 domain	Cytosol
LIG_SUMO_SIM_anti_2	STVLLVVQ TVLLVVQ	25–32 26–32	Motif for the antiparallel beta augmentation mode of non-covalent binding to SUMO protein	PML body, Nucleus, Nuclear body
LIG_TRAF2_1	PLEE	89–92	Major TRAF2-binding consensus motif	Cytosol
MOD_CK2_1	GRSSPYE LKQTVAE	44–50 53–59	CK2 phosphorylation site	Nucleus, Cytosol, Protein kinase CK2 complex
MOD_GSK3_1	YGLTKGCT	93–100	GSK3 phosphorylation recognition site	Cytosol, Nucleus
MOD_NEK2_1	LKQTVA	53–58	NEK2 phosphorylation motif with preferred Phe, Leu or Met	Centrosome, Ndc80 complex, Kinetochore, Cytosol, Nucleus
TRG_ER_diArg_1	LRLR	105–108	The di-Arg ER retention motif is defined by two consecutive arginine residues (RR) or with a single residue insertion (RXR)	Cytosol, Endoplasmic reticulum membrane, Integral protein

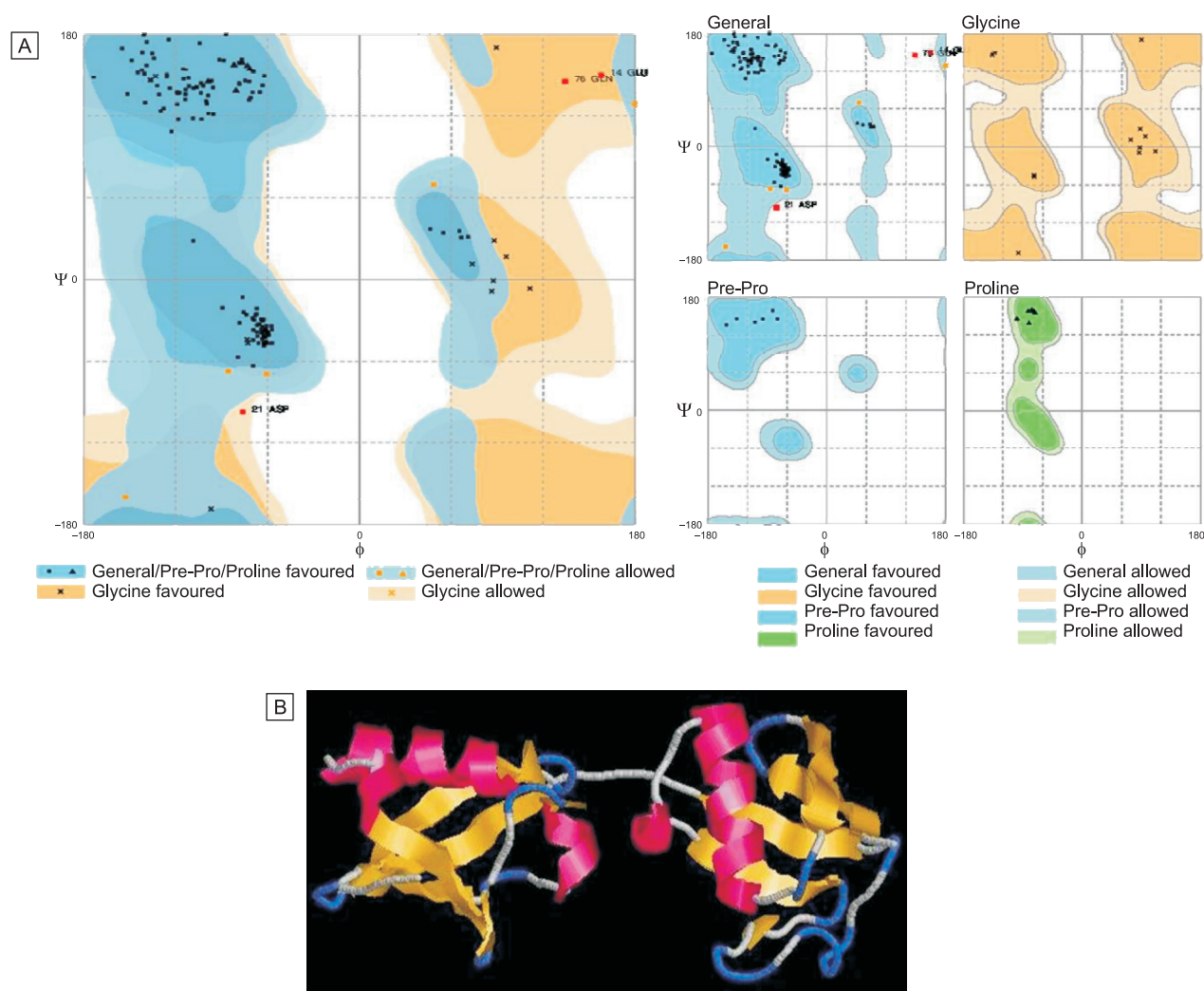


Fig. 2. A. Ramachandran plot of ISG15 protein generated using Modeller 9.8. B. Swiss pdb view of ISG15 protein model generated by Modeller 9.18 after energy minimization.

reactivity and deciphering the activity of protein (Kwok *et al.* 2002). In the present study, three dimensional structures were generated using modeller 9.18 and its ramachandran plot was plotted. BLASTP search was performed utilizing ISG15 protein sequence against PDB with default parameters to find suitable templates for homology modeling which can run as test sequence for model generation. Three test sequences [1Z2M_ACrystal Structure of ISG15, the Interferon-Induced Ubiquitin Cross Reactive Protein; 3SDL_C, Crystal structure of human ISG15 in complex with NS1 N-terminal region from influenza B virus; 5CHF_ACrystal structure of murine ISG15 in space group P21212] having maximum query coverage with high score and lower e-value was used as the template for homology modeling. The different pdb file generated of ISG15 protein was visualized in swiss pdb viewer (Fig. 2B). The sequence structure alignment used for model building was built by T-coffee. The model generated by swiss prot modeler gave only one model with GMQE value 0.81 and Q mean value as -1.34 (models not shown).

Ramachandran plot is the graphical way to identify energetically allowed regions for backbone dihedral angles

ϕ against ψ of amino acid residues in protein structure. It was found that 94.7%, residues fell in the most favored regions, 3.3% residues lies generously allowed regions and 2.0% residues lied in the disallowed conformations in ISG15 protein and was the most favored structure (Fig. 2A). The overall Verify 3D for the homology modeled structure was -0.02. This score indicates that the modeled structure of ISG15 protein is acceptable. For stabilization of any protein in 3D structure it requires proper folding which is determined by secondary structures present in it (Kwok *et al.* 2002). An evaluation of percentage of secondary structures revealed domination of alpha helix (31.17%) and random coil (32.47%) in ISG15 protein indicating folding energy for right handed is more favorable. Other secondary structure features predicted using self optimized prediction method was 11.69% beta strand and 24.68% extended strand. The graphical view of secondary structures by Phyre was given in Fig. 3.

The functional characterization of ISG15 protein was performed by identifying eukaryotic linear motifs (ELMs) defined as compact protein interaction sites composed of short stretches of adjacent amino acids (Table 1). This motif

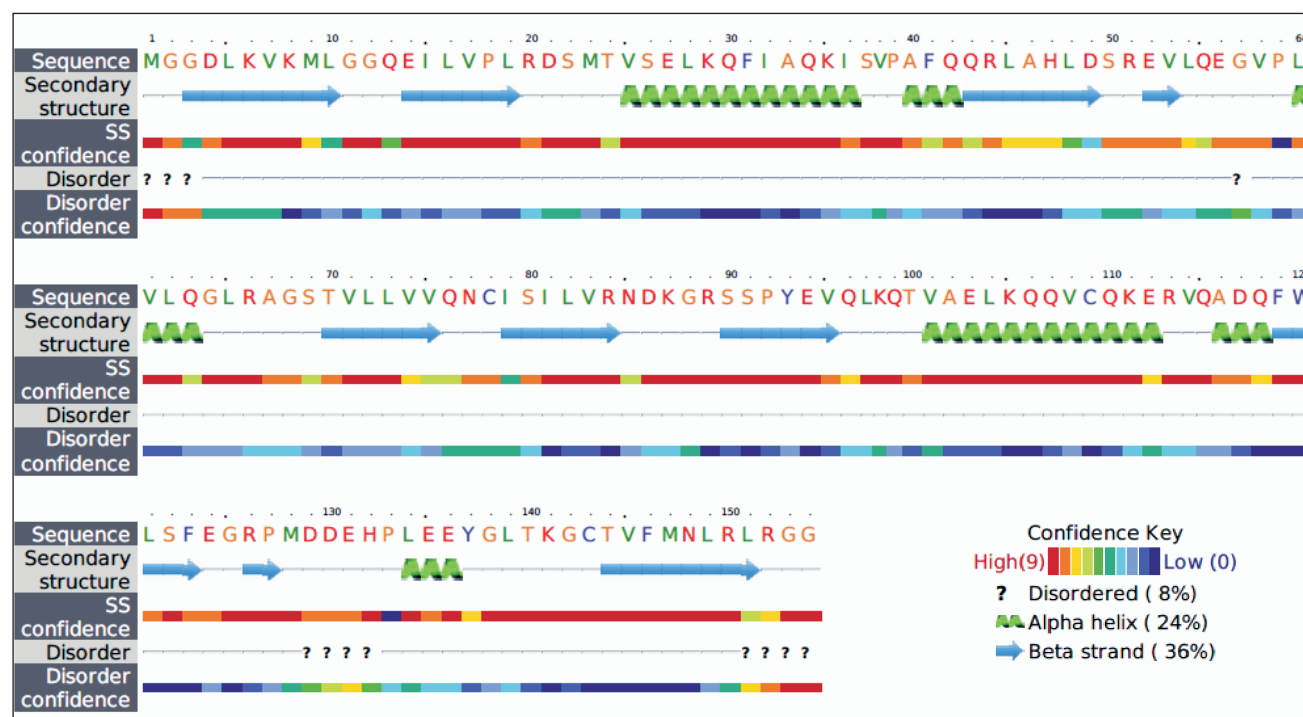


Fig. 3. Predicted secondary in ISG15 sequence using PHre2 software.

identification of different regions of ISG15 protein provided insights in identification of different docking parts involved in cell signaling process and may be predicted for its interaction with different molecules.

NetSurfP found that the ISG15 protein has a combination of buried and exposed amino acid residues which signifies the presence of transmembrane segments in this protein. The RSA (Relative Surface Accessibility) value ranges from 0.012 to 0.994. The detailed output of this prediction is not shown here (available with the authors).

There was strong correlation between stability of a protein and its physiochemical properties for predicting *in vivo* stability of a protein from its primary sequence (Guruprasad *et al.* 1990, Smialowski *et al.* 2007). The average molecular weight of ISG15 protein predicted was 17280.07 dalton. The total number of negatively charged residues present in ISG15 was (Asp + Glu) 18 and total number of positively charged residues (Arg + Lys) was 19. There were 2,458 number of atoms in ISG15 molecule having atomic composition as $C_{761}H_{1248}N_{216}O_{225}S_8$. Expasy's ProtParam computes the extinction coefficient (EC) of ISG15 for a 280 nm wavelength in water. Its EC value ranges 8605 Abs 0.1% (=1 g/l) 0.498, assuming all pairs of Cys residues form cystines to EC 8480 Abs 0.1% (=1 g/l) 0.491, assuming all Cys residues are reduced. The high EC value in ISG15 indicated the presence of high concentration of Cys, Trp and Tyr. Aliphatic index (AI) which was 99.87 for ISG15 determines the relative volume of an ISG15 protein occupied by aliphatic side chains and various predisposing factor responsible for the increase of thermal stability of protein. The instability index for given ISG15 protein was 49.77 which classified protein was

unstable in *in vitro* conditions indicating it was not suitable for *in vitro* studies which further diminishes its possibility of any heterologous expression. The half life of ISG15 protein indicates it is stable for 30 h (mammalian reticulocytes, *in vitro*), >20 h (yeast, *in vivo*) and >10 h (*E. coli*, *in vivo*). This predicted the time taken by ISG15 protein to disappear after its synthesis in the cell.

The negative GRAVY value, i.e. -0.196 of ISG15 protein consists of more hydrophilic residues which may be a clue towards its secretory nature and easily accessible to charged ions. Isoelectric point (pI) value of proteins is defined as the pH at which the net charge on the protein is zero. The computed isoelectric point (pI) value of ISG15 protein was 7.72 indicating that ISG15 was acidic in nature which is useful for wet lab extraction (through chromatographic methods). This also indicated that protein will be not soluble at these pI values and buffer pH required must be far away to achieve solubility of these proteins.

The ability to perform complex post-translational modifications is a major attribute that distinguishes mammalian cells from other production platforms used to produce recombinant proteins. Prediction of some Post-translational modifications (PTMs), such as glycosylation, methylation, phosphorylation and disulfide bond formation, was found to significantly worsen protein amenability to soluble expression and proper folding (Tokmakov *et al.* 2012). Importantly, many eukaryotic proteins require multiple PTMs to reach a native, biologically active conformation. PTMs can significantly change the integral characteristics of proteins that affect their stability and solubility, such as charge, hydrophobicity, solvent accessibility, etc. (Guruprasad *et al.* 1990, Karve and

Cheema, 2011). Thus, PTMs should also be considered as the major determinants of successful protein synthesis. Phosphorylation of proteins is involved in many signaling pathways function and it may lead to alteration in function of protein and its localization. There were 6 serine residues and 4 threonine residues in ISG15 protein having high threshold of binding phosphate group. Glycosylation plays a critical role in determining protein structure, function and stability. There were many asparagines, threonine and tryptophan residue in ISG15 protein but none of them have probability for having glycosylated. There was only one place for disulphide bond formation in ISG15 protein at 151 which was also a terminal region. Post-translational lysine methylation and acetylation were two major modifications found in ISG15 protein at lysine residues. It was found that there were 4 methylated lysine sites and 2 Acetyl lysine sites present in ISG15 protein sequence (Table 2).

Prokaryotic system has been used widely as expression system for recombinant protein production (Yin *et al.* 2007). Low phosphorylation, least disulphide bond formation, low methylation and non glycosylation in ISG15 protein provide a great advantage for expression in wide heterologous expression systems (Sahadev *et al.* 2008). The domain identified in ISG15 protein was Ubiquitin-like proteins. Ubiquitin-mediated proteolysis is part of the regulated turnover of proteins required for controlling cell cycle progression. Ubiquitination usually results in a covalent bond between the C-terminus of ubiquitin and the epsilon-amino group of a substrate lysine. This plays important role in binding of ubiquitin with ISG15 protein in response to release of Interferon-tau from implanting foetus.

A suitable peptide having more immunogenicity in different expression system is the foremost requirement for ISG15 protein to act as biomarker. Immunological characterization has a key role in designing of various antibody based detection kits for early pregnancy diagnosis which are immediately required in buffalo (Green *et al.* 2005). Therefore, present study revealed 9 immunogenic B cell epitope in ISG15 protein with different peaks

Table 2. Different acetylation and methylation sites in ISG15 protein

Protein name	Position of site	Flanking residues	Predicted result	SVM Probability
DLKVKM	43	ILVRND-K-GRSSPY	Acetyl lysine	0.540954
DLKVKM	54	PYEVQL-K-QTVAEL	Methylated lysine	0.677102
DLKVKM	61	QTVAEL-K-QQVCQK	Methylated lysine	0.593230
DLKVKM	67	KQQVCQ-K-ERVQAD	Methylated lysine	0.511482
DLKVKM	97	EEYGLT-K-GCTVFM	Methylated lysine	0.644464
DLKVKM	97	EEYGLT-K-GCTVFM	Acetyl lysine	0.500000

indicating antigenicity of different regions and providing strong indication of antigenic nature of these proteins (Table 3). These linear epitopic sites can be targeted for production of antibody in *in-vivo* system or recombinant antibody can be prepared against these antigenic sites which are currently required for detection of this protein in serum of pregnant buffalo at 18 to 21 day after artificial insemination. This protein can be a strong indicator for predicting embryonic mortality in animals (Sheikh *et al.* 2018). Therefore, any assay quantifying the increased expression of ISG15 can act as biomarker for early pregnancy diagnosis as well as conception failure/or embryonic mortality.

Table 3. Different biepiped linear epitope predicted in ISG15 protein

Start	End	Peptide	Length
1	3	MGG	3
22	25	SMTV	4
38	41	VPAF	4
52	55	EVLQ	4
66	67	RA	2
86	97	DKGRSSPYEVQL	12
102	103	AE	2
108	116	VCQKERVQA	9
125	140	GRPMDDEHPLEEYGLT	16

The present study concludes that ISG15 gene is highly expressed at 18 to 21 days of pregnancy of animal after artificial insemination. The phylogenetic studies reveals its high similarity with other related species like cattle indicating its cross reactivity. The prediction of various properties like structural, physiochemical, functional and immunological are required for utilization of ISG15 protein in any heterologous expression system. These properties provide good indication towards favorable characteristics of this molecule like more favoured structure, less post translation modifications (Glycosylation, phosphorylation, disulphide bridges) and number of antigenic epitopes make it highly suitable marker molecule which can be employed in early pregnancy diagnostic kits in buffalo.

ACKNOWLEDGEMENTS

The authors would like to thank veterinarians and non teaching staffs for collection of samples. They also acknowledge funding support from Council of Scientific and Industrial Research (CSIR), New Delhi for providing fellowship assistance and Rashtriya Krishi Vikas Yojana (RKVY) for the research cost of whole experiment.

REFERENCES

- Arnold K L, Bordoli J K and Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195–201.
- Austin K J, Ward S K, Teixeira M G, Dean V C, Moore D W and Hansen T R. 1996. Ubiquitin cross-reactive protein is released by the bovine uterus in response to interferon during early

- pregnancy. *Biology of Reproduction* **54**: 600–606.
- Austin K J, Bany B M, Belden E L, Rempel L A, Cross J C and Hansen T R. 2003. Interferon stimulated gene-15 (ISG15) expression is up-regulated in the mouse uterus in response to the implanting conceptus. *Endocrinology* **144**: 3107–13.
- Balhara A K, Gupta M, Mohanty A K, Phulia S K, Sharma R K, Singh S and Singh I. 2014. Changes in sera proteome in relation to day of pregnancy in early pregnant buffaloes. *Indian Journal of Animal Sciences* **84**(4): 400–409.
- Batra K. 2017. 'In vitro expression, molecular and immunological characterization of early pregnancy related interferon tau stimulated gene of *Bubalus bubalis*.' PhD Thesis, LUVAS, Hisar, Haryana, India.
- Batra K, Nanda T, Kumar A, Kumar V, Gopal G J and Maan S. 2018a. Molecular cloning and expression kinetics of serum Interferon stimulated gene for early pregnancy detection. *Indian Journal of Animal Research*. DOI: 10.18805/ijar.B-3636.
- Batra K, Nanda T, Kumar A, Gupta A K, Kumari R, Kumar V, Sheoran N and Maan S. 2018b. Molecular cloning and characterization of Mx2 for early pregnancy diagnosis in *Bubalus bubalis*. *Proceedings of the National Academy of Sciences India Section B Biological Sciences*. <https://doi.org/10.1007/s40011-018-0993-x>
- Batra K, Kumar A, Maan S, Kumar V, Kumari R and Nanda T. 2018c. Recombinant interferon stimulated protein 15 (rISG15) as a molecular marker for detection of early pregnancy in *Bubalus bubalis*. *Animal Reproduction Science* **197**: 106–16.
- Bazer F W, Spencer T E, Johnson G A, Burghardt R C and Wu G. 2009. Comparative aspects of implantation. *Reproduction* **138**: 195–209.
- Bebington C, Doherty F J and Fleming S D. 1999. Ubiquitin cross-reactive protein gene expression is increased in decidualized endometrial stromal cells at the initiation of pregnancy. *Molecular Human Reproduction* **5**: 966–72.
- Blom N, Gammeltoft S and Brunak S. 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology* **294**(5): 1351–62.
- Chauhan J S, Rao A and Raghava G P S. 2013. In silico Platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences. *PloS one* **8**(6):e67008.
- Dinkel H, Van K R, Michael S, Kumar M, Uyar B, Altenberg B, Milchevskaya V, Schneider M, Kuhn H, Damerell V, Diebel S, Kalman S, Klein S, Knudsen A C, Mader C, Merrill S, Staudt A, Thiel V, Welti L, Davey NE, Diella F and Gibson T J. 2016. ELM 2016 –data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Research* **44**(4): 294–300.
- Forde N, Carter F, Spencer T E, Bazer F W, Sandra O, Mansouri-Attia N, Okumu L A, McGettigan P A, Mehta J P, McBride R, O'Gaora P, Roche J F and Lonergan P. 2011. Conceptus-induced changes in the endometrial transcriptome: how soon does the cow know she is pregnant?. *Biology of Reproduction* **85**: 144–56.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M R and Appel R D. 2005. Protein Identification and Analysis Tools on the ExPASy Server, pp 571–607. *The proteomics protocols Handbook*. (Eds) John M. Walker. Humana Press.
- Gill S C and Von Hippel P H. 1989. Extinction coefficient. *Analytical Biochemistry* **182**: 319–28.
- Green J A, Parks T E, Avalle M P, Telugu B P, McLain A L, Peterson A J, McMillan W, Mathialagan N, Hook R R, Xie S and Roberts R M. 2005. The establishment of an ELISA for the detection of pregnancy-associated glycoproteins (PAGs) in the serum of pregnant cows and heifers. *Theriogenology* **63**(5): 1481–1503.
- Green J C, Okamura C S, Poock S E and Lucy M C. 2010. Measurement of interferon-tau (IFN-tau) stimulated gene expression in blood leukocytes for pregnancy diagnosis within 18–20 d after insemination in dairy cattle. *Animal Reproduction Science* **121**: 24–33.
- Guex N and Peitsch M C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714–23.
- Guruprasad K, Reddy B and Pandit M W. 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* **4**:155–164.
- Hamerman J A, Hayashi F, Schroeder L A, Gygi S P, Haas A L, Hampson L, Coughlin P, Aebersold R and Aderem A. 2002. Serpin 2a is induced in activated macrophages and conjugates to a ubiquitin homolog. *Journal of Immunology* **168**: 2415–23.
- Han H, Austin K J, Rempel L A and Hansen T R. 2006. Low blood ISG15 mRNA and progesterone levels are predictive of non-pregnant dairy cows. *Journal of Endocrinology* **191**: 505–12.
- Hansen T R, Austin K J and Johnson G A. 1997. Transient ubiquitin cross-reactive protein gene expression in the bovine endometrium. *Endocrinology* **138**: 5079–82.
- Hansen T R, Austin K J, Perry D J, Pru J K, Teixeira M G and Johnson G A. 1999. Mechanism of action of interferon-tau in the uterus during early pregnancy. *Journal of Reproduction Fertility* **54**: 329–39.
- Ikai A J. 1980. Thermo stability and aliphatic index of globular proteins. *Journal of Biochemistry* **88**: 1895–98.
- Jain A, Baviskar P S, Kandasamy S, Kumar R, Singh R, Kumar S, Agarwal S K, Joshi P and Mitra A. 2012. Interferon stimulated gene 15 (ISG15): molecular characterization and expression profile in endometrium of buffalo (*Bubalus bubalis*). *Animal Reproduction Science* **133**(3–4): 159–68.
- Johnson G A, Austin K J, Van Kirk E A and Hansen T R. 1998. Pregnancy and interferon-tau induce conjugation of bovine ubiquitin cross-reactive protein to cytosolic uterine proteins. *Biology of Reproduction* **58**: 898–904.
- Johnson G A, Spencer T E, Hansen T R, Austin K J, Burghardt R C and Bazer F W. 1999. Expression of the interferon tau inducible ubiquitin crossreactive protein in the ovine uterus. *Biology of Reproduction* **61**: 312–18.
- Joyce M M, Hansen T R and Johnson G A. 2002. Interferon-stimulated gene 17 is expressed in the porcine uterus and may be critical to placental development across species. *Biology of Reproduction* **66**: 185.
- Joyce M M, White F J, Burghardt R C, Muniz J J, Spencer T E, Bazer F W and Johnson G A. 2005. Interferon stimulated gene 15 conjugates to endometrial cytosolic proteins and is expressed at the uterineplacental interface throughout pregnancy in sheep. *Endocrinology* **146**: 675–84.
- Karve T M and Cheema K A. 2011. Small changes huge impact: The role of protein posttranslational modifications in cellular homeostasis and disease. *Journal of Amino Acids*: 207691.
- Kelley L A, Mezulis S, Yates C M, Wass M N and Sternberg M J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**(6): 845–58.
- Kizaki K, Shichijo-Kizaki A, Furusawa T, Takahashi T, Hosoe M

- and Hashizume K. 2013. Differential neutrophil gene expression in early bovine pregnancy. *Reproductive Biology and Endocrinology* **11**: 6.
- Kwok S C, Mant C T and Hodges R S. 2002. Importance of secondary structural specificity determinants in protein folding: Insertion of a native β -sheet sequence into an α -helical coiled-coil. *Protein Science* **1**(6): 1519–31.
- Kyte J and Doolittle R F. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**: 105–32.
- Larsen J E, Lund O and Nielsen M. 2006. Bepipred Linear Epitope Prediction: Improved method for predicting linear B-cell epitopes. *Immunome Research* **2**: 2.
- Loeb K R and Haas A L. 1992. The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. *Journal of Biological Chemistry* **267**: 7806–13.
- Lovell S C, Davis I W, Arendall W B, de Bakker P I W, Word J M, Prisant M G, Richardson J S and Richardson D C. 2002. Structure validation by C α geometry: ϕ , ψ and C β deviation. *Proteins: Structure Function and Genetics* **50**: 437–50.
- Luthy R, Bowie J U and Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature* **356**: 83–85.
- Malakhov M P, Kim K I, Malakhova O A, Jacobs B S, Borden E C and Zhang D E. 2003. High-throughout immunoblotting. Ubiquitin-like protein ISG15 modifies key regulators of signal transduction. *Journal of Biological Chemistry* **278**: 16608–13.
- Matsuyama S, Kojima T, Kato S and Kimura K. 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. *Reproductive Biology and Endocrinology* **10**: 21.
- Meyerholz M M, Mense K, Knaack H, Sandra O and Schmicke M. 2016. Pregnancy-Induced ISG-15 and MX-1 Gene Expression is Detected in the Liver of Holstein-Friesian Heifers During Late Peri-Implantation Period. *Reproduction in Domestic Animals* **51**(1): 175–77.
- Notredame C, Higgins D G and Heringa J. 2011. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**(1): 205–17.
- Perry D J, Austin K J and Hansen T R. 1999. Cloning of interferon-stimulated gene 17: the promoter and nuclear proteins that regulate transcription. *Molecular Endocrinology* **13**: 1197–1206.
- Petersen B, Petersen T N, Andersen P, Nielsen M and Lundegaard C. 2009. A generic method for assignment of reliability scores applied to solvent accessibility predictions. *BMC Structural Biology* **9**: 51.
- Ramachandran G N, Ramakrishnan C and Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **7**: 95–99.
- Rempel L A, Francis B R, Austin K J and Hansen T R. 2005. Isolation and sequence of an interferon- δ inducible, pregnancy and bovine interferon stimulated gene product 15 (ISG15)-specific, bovine ubiquitin activating E1-like (UBE1L) enzyme. *Biology of Reproduction* **72**: 365–72.
- Rodriguez M R, Monte K, Thackray L B and Lenschow D J. 2014. ISG15 Functions as an Interferon-Mediated Antiviral Effector Early in the Murine Norovirus Life Cycle. *Journal of Virology* **88**(16): 9277–86.
- Sahdev S, Khattar S K and Saini K S. 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry* **307**: 249–264.
- Sali A and Blundell T L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**: 779–815.
- Sapay N, Guermeur Y and Deleage G. 2006. Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier. *BMC Bioinformatics* **7**(1):255.
- Sheikh A A, Hooda O K, Kalyan A, Kamboj A, Mohammed S, Alhussien M, Reddi S, Shimray P G, Rautela A, Pandita S, Kapila S, De S and Dang A K. 2018. Interferon-tau stimulated gene expression: A proxy to predict embryonic mortality in dairy cows. *Theriogenology* **120**: 61–67.
- Shi S P, Qiu J D, Sun X Y, Suo S B, Huang S Y and Liang R P. 2012. PLMLA: prediction of lysine methylation and lysine acetylation by combining multiple feature. *Molecular BioSystems* **8**(5): 1520–27.
- Singh S K, Agarwal S K, Shankar U and Gupta L K. 2005. Purification and characterization of protein(s) from placental extracts in buffalo. *Indian Journal of Animal Sciences* **75** (7):769–800.
- Smialowski P, Martin-Galiano A J, Mikolajka A, Girschick T, Holak T A and Frishman D. 2007. Protein solubility: sequence based prediction and experimental verification. *Bioinformatics* **23**(19): 2536–42.
- Spencer T E, Sandra O and Wolf E. 2008. Genes involved in conceptus endometrial interactions in ruminants: insights from reductionism and thoughts on holistic approaches. *Reproduction* **135**: 165–79.
- Soumya N, Das D, Jeyakumar S, Mondal S, Mor A and Mundhe U. 2016. Differential expression of ISG 15 mRNA in peripheral blood mononuclear cells of nulliparous and multiparous pregnant versus non-pregnant Bos indicus cattle. *Reproduction in Domestic Animals* **52**: 97–106.
- Thakur N, Singh G, Paul A, Bharati J, Rajesh G, Gm V, Chouhan V S, Bhure S K, Maurya V P, Singh G and Sarkar M. 2017. Expression and molecular cloning of interferon stimulated genes in buffalo (*Bubalus bubalis*). *Theriogenology* **100**: 50–58.
- Thatcher W W, Guzeloglu A, Mattos R, Binelli M, Hansen T R and Pru J K. 2001. Uterine-conceptus interactions and reproductive failure in cattle. *Theriogenology* **56**: 1435–50.
- Tokmakov A A, Kurotani A, Takagi T, Toyama M, Shirouzu M, Fukami Y and Yokoyama S. 2012. Multiple post-translational modifications affect heterologous protein synthesis. *Journal of Biological Chemistry* **287**(32): 27106–16.
- Yin J, Li G, Ren X and Herrler G. 2007. Select what you need; a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *Journal of Biotechnology* **127**: 335–47.