Indian Journal of Biotechnology
Vol. 18, April 2019, pp 97-107

Molecular characterization of OAS1 as a biomarker molecule for early pregnancy diagnosis in *Bubalus bubalis*

Kanisht Batra*, Trilok Nanda, Aman kumar, Rajni Kumari, Vinay Kumar and Sushila maan
Department of Animal Biotechnology, Lala Lajpat Rai University of Veterinary and Animal Sciences, Hisar, Haryana 125004, India

Received 7 July 2017; revised 26 March 2019; accepted 30 March 2019

OAS1 (2′-5′-oligoadenylate synthetase) is a type I interferon molecule which is responsible for antiviral activity and known for its action in innate immunity. It was also reported to be stimulated by the release of interferon tau (IFN-t) in the serum of pregnant animals during conceptus implantation. OAS1 is studied in most of the ruminants except for buffalo (*Bubalus bubalis*), which is well known for high milk production. Therefore, in the present study OAS1 mRNA was isolated from pregnant buffalo (*B. bubalis*) at 18-21 day of pregnancy. The ORF region of OAS1 was amplified and sequenced for characterization of this protein. *In silico* analysis of protein was done in order to find the suitability of protein as a candidate molecule for early pregnancy diagnostic kits. The functional characterization identifies various motifs present in the protein which are responsible for its interaction with other proteins. Physiochemical properties predict the protein nature during *in vitro* conditions. Further, immunogenic studies revealed OAS1 is highly antigenic nature and its suitability for use *in vivo* conditions. In conclusion, OAS1 protein expression in buffalo is a good indicator of conceptus implantation and has suitable properties for being used in early and precise pregnancy diagnostic kits.

Keywords: OAS1, buffalo, expression, molecular characterization

Introduction

Reproductive biotechnology in domestic buffalo has not been developed as to the extent it has in cattle due to its more distribution in developing countries like India. Although artificial insemination has been widely practiced which have improved reproductive status of the animals, still it is suffering from limitations related to puberty, calving seasonality, long postpartum anoestrus, low conception rate, poor expression of estrus signs and prolonged calving intervals[1,2]. One of the best possible ways to increase reproductive efficiency is to develop the tool which can detect the pregnancy early and precisely. Non-return to estrus after 21 days and rectal palpation of the pregnant uterus have been the most common methods adopted for pregnancy diagnosis in ruminants. However, estrus behavior in buffalo has a lower intensity than other species resulting in absence of behavioural signs of ovulation. Even palpation of the amniotic vesicles per-rectally and chorio-allantoic membranes slipping between the thumb and forefinger can be misleading for pregnancy detection. Therefore, a marker molecule which can detect pregnancy at the earliest possible time is the foremost requirement to improve the reproductive efficiency in buffalo.

Interferon-tau (IFN-t), a novel type I interferon is first produced by the conceptus between days 14–16 after conception in cattle[3]. This is well characterized protein produced by the ruminant trophoblast during initiation of pregnancy at 14th day. IFN-t plays an important role by preventing increases in endometrial estrogen and oxytocin receptors, to dissolve the oxytocin-induced luteolytic pulses of PGF2 alpha and maintains corpus luteum (CL) function[4]. Some IFN-t escapes the uterus and leaks out which can be detected in blood[5,6]. Leukocytes respond to IFN-t by expressing interferon-stimulated genes (ISG) as interferon-stimulated protein such as 15 kDa (ISG15), myxovirus resistance 2 (Mx2) and 2′-5′-oligoadenylate synthetase (OAS1), in peripheral blood leukocytes during early pregnancy. The utility of ISGs for the purpose of pregnancy detection in dairy cattle has been demonstrated recently[7-10]. Out of these proteins, OAS 1 is one of the candidate molecules which can be used for detection of early pregnancy in animals.

The interferon-stimulated genes are also activated in response to viral infection and play an important role in the host immunity. These type I IFNs induce several proteins with antiviral activity such as double-stranded RNA-activated protein kinase (PKR),

*Author for correspondence:
drkanishtbatra@gmail.com

2′-5′-oligoadenylate synthetases (OASs), RNase L, and the Mx protein GTPases. OAS1 is an enzyme which is responsible for catalyzing the synthesis of oligoadenylates of the general structure ppp(A2′p)$_n$A, commonly abbreviated as 2-5A. It is responsible for inhibition of protein synthesis in viral replication and in some cases it causes apoptosis.

This molecule was first reported with a key role in pregnancy detection with increased level of expression on 18th day of pregnancy in heifers (approximately 1 year of age)[11]. Microarray analysis has also revealed that 17 genes including OAS1 had a greater than or equal to 2-fold change in gene expression on 18th day compared with 15th day of pregnancy[8]. The present study was planned to find out the suitability of OAS1 molecules for the development of early pregnancy diagnostic kit. Therefore, in the present study, OAS1 was isolated from *B. bubalis* and reverse genetics approach was applied to reveals it's *in silico* characteristics.

Materials and Methods

Collection of Sample

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. Fifty Murrah buffaloes were selected from an organized farm of LUVAS, Hisar. Animal showing estrus signs were inseminated and used for collection of blood sample. The blood samples were collected using EDTA as anticoagulant by jugular vein puncture under sterile conditions.

RNA Isolation

Blood samples were immediately processed for extraction of total RNA using TRIzol Reagent (Ambion) with minor modifications. Briefly, 600 µl of blood was mixed with 800 µl of TRIzol reagent and vortexed uniformly. Chloroform was added for separation of organic phase and aqueous phase containing RNA. The RNA in aqueous phase was precipitated by adding 500 µl isopropanol and incubation overnight at -20°C. The RNA pellet was centrifuged at 12000 rpm for 20 min and was washed with 1 ml of 70% ethanol and resuspended in 20 µl of nuclease free water. RNA isolated was immediately reverse transcribed and was stored at -20ºC till used.

Reverse Transcription

It was carried out from the total RNA isolated by TRIzol reagent using Thermo Scientific RevertAid Strand cDNA Synthesis Kit as per manufacturer's first instruction. Briefly, NFW (5.00 µL), 5X RT buffer (4 µL), 10 mM dNTPs (Fermentas) (2 µL), total RNA (5 µL), RT 200 IU/µL (1 µL), random hexamer (2 µL). The cDNA prepared was stored at –20˚C till further use.

Primer Designing

Primers were designed using Primer BLAST software from NCBI using the gene sequence of accession no. NM_001040606.1 and XM_006052997.1 (*B. bubalis*). These primers were designed to target the complete ORF region of OAS1 gene. Primers sequences are forward primer 5' CACCATGGAGCTCAGAAATACCCC 3' and reverse primer 5' TCAGAGGATGGTACATGTCCAG 3'.

Conventional PCR

Primer pair flanking starting region of ORF was used for amplification of full length OAS1 ORF region of expected size 1170 bp. PCR was performed in thermal cycler (Veriti Applied Biosystem) in 25 µl reaction containing 5 µl of template cDNA using high fidelity Fusion *Taq* mastermix (2X concentration) with GC buffer (NEB). Cyclic conditions were optimized for PCR 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.

Sequencing of PCR Products

Purified PCR products were sequenced using BigDye terminator v3.1 cycle sequencing kit from 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 phylogenetic tree was constructed to show relatedness between all available OAS1 sequences in different species. This sequence was translated into amino acid sequences using expasy translate tool and then further characterized.

Prediction of Primary Structure

The OAS1 protein primary structure was predicted using academic version of MODELER9.8[12]. The 3D model of OAS1 protein is finally generated by optimization of the molecular pdf with little violation of restrains constraint. Sequence alignment between buffalo OAS1 protein and other proteins available was done using T Coffee program[13]. Model was also generated using automated Swiss Prot modeller[14]. OAS1 models were predicted sequentially with the global model quality estimate (GMQE) and the QMEAN score. The structure was confirmed using

Procheck and VERIFY3D[15]. Ramachandran plot was generated to check the stereo chemical properties of OAS1 protein structure which has been modeled[16].

Prediction of Secondary Structure

The secondary structure of OAS1 protein was predicted using SOPMA (Self-Optimized Prediction Method with Alignment) tool[17].

Prediction of Linear Motif and Functional Characterization

The Eukaryotic Linear Motiff (ELM)[18] prediction tools was used which scans the submitted OAS1 protein sequences for matches with the regular expressions defined in ELM.

Physiochemical Properties of Protein

The physicochemical properties were calculated using Expasy's Prot-Param[19] prediction server. This predicts the different properties of OAS1 protein such as molecular weight, theoretical isoelectric point (pI), total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY).

Prediction of Accessible Surface Area (ASA)

The accessible surface area of the OAS1 protein was predicted through NetSurfP[20] server of ExPaSy software.

Predicting Phosphorylation Sites, Glycosylation Sites, Methylation and Acetylation and Sumoylation Sites

NetPhos server is a tool to predict different sites in OAS1 protein which can be phosphorylated[21]. The glycosylation site was predicted using GLCO ep software[22]. The fourth-generation GPS algorithm integrated with the Particle Swarm Optimization (PSO)[23] method was employed for prediction of sumoylation sites in OAS1 protein. Prediction of potential methylation and acetylation of OAS1 protein sequence was done using *in silico* tool PLMLA (Prediction of Potential Lysine Methylation and Lysine Acetylation)[24].

Prediction of Disulfide Bridges

CYS_REC (http://www.softberry.com/cgi-bin/programs/propt/cys_rec.) tool 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 OAS1 protein sequence.

Conserved Domain Identification for Function Prediction

The OAS1 protein sequence was used as input using NCBI CDD-BLAST tool (available at *http://www.ncbi.nlm.nih.*) for searching conserved domains.

Prediction of Immunogenic Sites

The B cell linear epitopes were predicted using IEDB analysis software[25] and conformational epitopes using the OAS1 sequence[26].

Results

Amplification of OAS1 Product

ORF region of OAS1 was amplified resulting in product size of 1170 bp (Fig. 1). The contig sequence obtained was submitted in gene bank with accession number MF402915 and further aligned with the available sequences in the GenBank using MEGA 6.0 programme. The Phylogenetic analysis of ORF sequence revealed 99% sequence identity with predicted *B. bubalis* OAS1 isoform, 97% with *Bos indicus*, 94% with sheep and goat sequence (Fig. 2).

Model Building

BLASTp search was performed using OAS1 sequence against PDB with default parameters to find suitable templates for homology modeling which can run as test sequence for model generation. Three test sequence [4RWN_A, Chain A, crystal structure of the pre-reactive state of porcine 4IG8_A, Chain A, structural basis for cytosolic double-stranded RNA surveillance by human 1PX5_A, Chain A, crystal structure of the 2'-specific and double-stranded RNA-activated interferon-induced antiviral protein 2'-5'-oligoadenylate Synthetase] having maximum identity with high score and lower e-value were used as the template for homology modeling. The pdb file generated was visualized in swiss pdb viewer as shown in Figure 3B. The model generated by swiss prot modeler gave only one model with GMQE value 0.78 and Q mean value as -0.30. Sequence alignment between buffalo OAS1

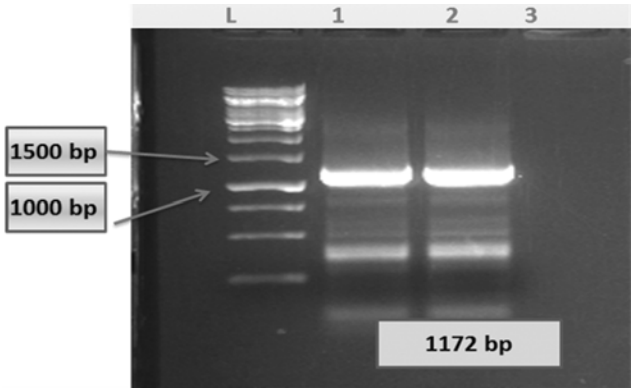

Fig. 1 — Agarose gel electrophoresis of PCR products generated using ORF specific OAS1 primer pair. Lane 1-2: Field samples at 18 day of pregnancy.

Fig. 2 — Unrooted neighbor joining phylogenetic tree drawn from complete ORF sequences of OAS1.

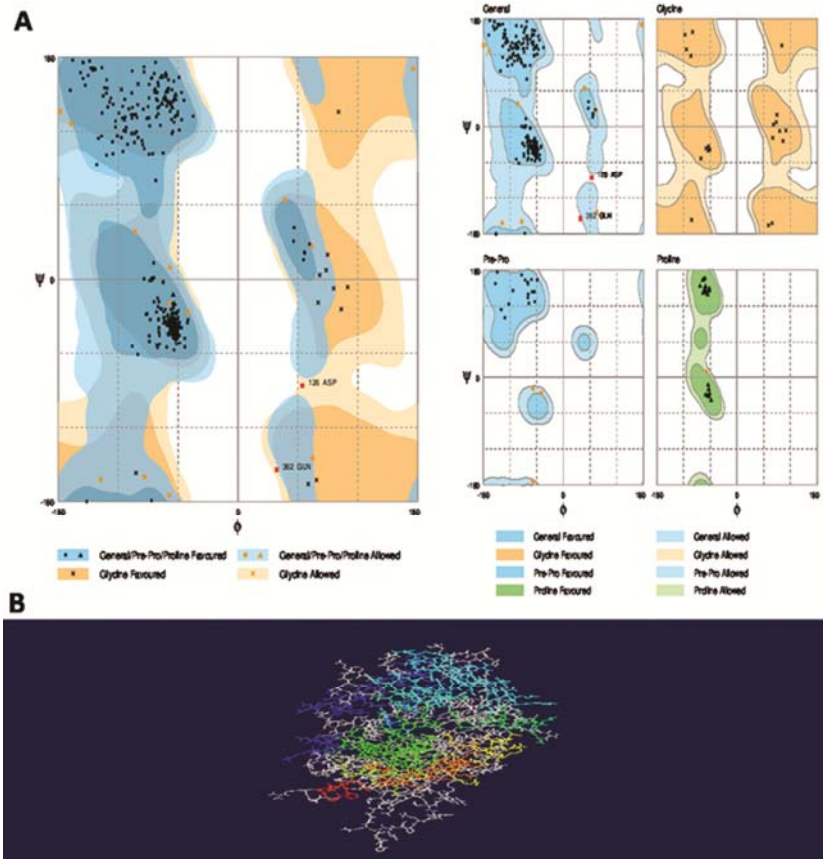

Fig. 3A — Ramachandran plot of OAS1 generated using Modeller 9.8 and 3B-Swiss pdb view of OAS1 model generated by Modeller 9.18 after energy minimization.

protein and other proteins was done using T Coffee program.

Protein Structure Validation

To validate the homology modeled buffalo OAS1 protein structure, a Ramachandran plot was drawn from the rampage site and the structure was analyzed by Procheck and verify 3D, a well-known protein structure checking program. It was found that the phi/psi angles of 96.1% residues fell in the most favored regions, 3.4% residues lied in the allowed

regions and 0.5% residues lied in the disallowed conformations (Fig. 3A). The overall Verify 3D for the homology modeled structure was -0.02. This score indicates that the modeled structure of OAS1 protein is acceptable.

Secondary Structure

The secondary structure features were predicted using self optimized prediction method and represented in Table 1. These results revealed that alpha helix was dominated among secondary structure elements followed by random coils, extended strand and beta turns.

Motiff Identification

ELM (Eukaryotic Linear Motiff) found that there are 41 different types of motifs recognized by different amino acid sequence in protein are given in

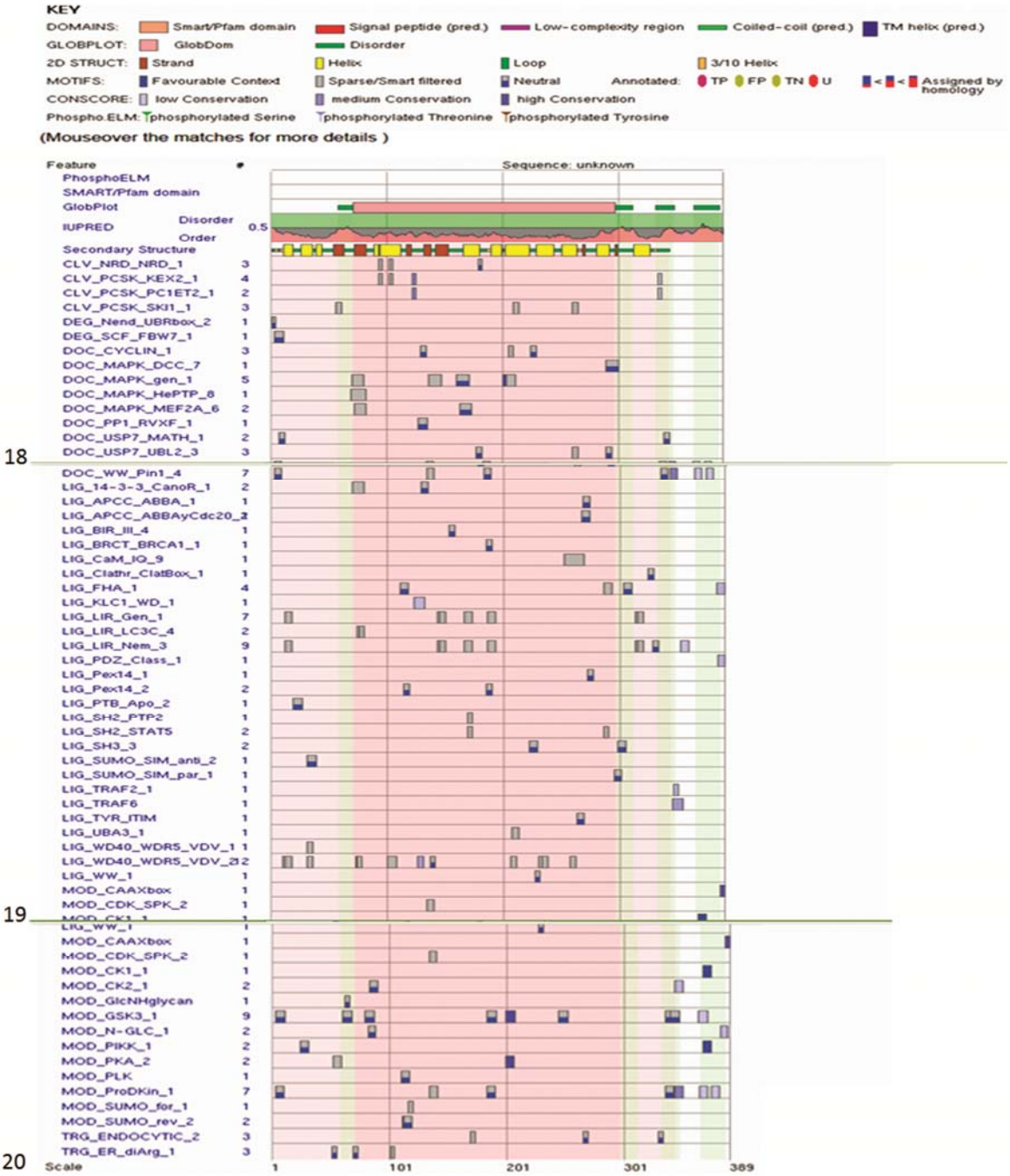

Fig. 4 — Graph depicting various ELM motiffs present in the OAS1 sequence.

Figure 4. These motif functionally characterize the major portion of OAS1 protein is involved in cell signaling processes and have various kinase and phosphorylation sites.

Accessible Surface Area

NetSurfP found that the hypothetical 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 0.065 to 0.600.

Physiochemical Properties of Protein

Various physiochemical properties of protein of OAS1 are given in Table 2. The EC ranges was 64120 Abs 0.1% (=1 g/l) 1.419, assuming all pairs of cystein residues form cystines to Ext. coefficient 63370 Abs 0.1% (=1 g/l) 1.402, assuming all cystein residues are reduced. The high EC value of indicated the presence of high concentration of cystein, tryptophan and tyrosine.

Phosphorylation and Sumoylation Sites in OAS1

NetPhos server was used to predict phosphorylation site at threonine, serine and tyrosine in OAS1 sequence**.** Various sites of phosphorylation and sumoylation sites in OAS1 were shown in Table 3. It was found that there were 21 serine, 20 threonine and 13 tyrosine but only 19 serine residues, 6 threonine residues and 2 tyrosine were above threshold levels which can be predicted as phosphorylated. There were three sumolytation sites at 57, 79 and 274 position in OAS1 sequence.

Methylation and Acetylation Sites

There were 11 methylated lysine sites in OAS1 sequence and 8 acetyl lysine sites present in OAS1 sequence. The position and probability of these sites is given in Table 4.

Glycosylation Sites

OAS1 sequences predicted the possibility of glycosylation at different sites. It was found that there were one potential N linked glycosylation sites and 4 O linked glycosylations site only in the OAS1 sequence in Table 5.

Disulphide Bridges in Protein

CYS_REC identifies the positions of various cysteines in OAS1 sequence, the total number of cysteines present and computed the most probable S-S bond pattern of the pairs in a protein sequence. Twelve cysteines are found at positions: 37, 44, 47, 108, 176, 188, 218, 244, 269, 314, 331 and 386. The

Table 1 — Different percentage of secondary structures given by SOPMA for OAS1

Alpha helix	45.76%
3_{10} helix	0.00%
Pi helix	0.00%
Beta bridge	0.00%
Extended strand	12.85%
Beta turn	7.20%
Bend region	0.00%
Random coil	34.19%
Ambiguous state	0.00%
Other states	0.00%

Table 2 — Different physiochemical properties of protein of OAS1

Molecular weight (Dalton)	45187.66
Isoelectric point (pI)	8. 34
Negatively charged residues (Asp + Glu)	53
Positively charged residues (Arg + Lys)	57
Number of atoms	6350
EC ranges Abs 0.1% (=1 g/l)	1.419-1.402
Instability index	46.56
Aliphatic index (AI)	78.41
GRAVY value	0.567

Table 3 — Position and function of phosphorylation and sumoylation sites in OAS1

Amino acid	Sequence	Position	Score	Kinase
T	ELRNTPAGS	6	0.929	unsp
S	TPAGSLDKF	10	0.538	cdc2
S	RVRVSKVVK	55	0.809	PKC
S	VKGGSSGKG	62	0.577	PKC
S	KGGSSGKGT	63	0.992	unsp
T	GKGTTLRGR	68	0.912	unsp
S	LRGRSDADL	73	0.997	unsp
S	TNLTSFREQ	86	0.870	unsp
T	QREETFEVK	113	0.792	unsp
S	FVLRSPKLD	137	0.870	unsp
T	LGQLTKGYR	159	0.581	PKC
S	YRPDSRVYV	166	0.692	unsp
S	EGEFSPCFT	186	0.991	unsp
T	KNRPTKLKS	202	0.708	PKC
S	TKLKSLIRL	206	0.906	unsp
T	QGFQTVLKL	257	0.603	PKC
Y	WEKNYNSEN	276	0.572	unsp
S	KNYNSENPI	278	0.523	CKII
S	KDTYSWERL	315	0.992	unsp
S	KRDGSPVGS	338	0.994	unsp
S	SPVGSWDVS	342	0.659	unsp
S	SWDVSVRPS	346	0.938	unsp
S	PQEHSDLMY	351	0.537	CKII
Y	SDLMYEAYD	355	0.809	unsp
S	QHYRSSPGT	366	0.914	unsp
S	HYRSSPGTQ	367	0.937	unsp
S	HGGASPQVE	377	0.686	unsp
K	REETFEVKFEVQKRQ	73	11.449	Sumoylation
K	SFVLRSPKLDQEVEF	95	3.311	Sumoylation
K	LDYPCFKKRDGSPVG	290	3.761	Sumoylation

S = Serine, Y = Tyrosine, K = Lysine

Table 4 — Different acetylation and methylation sites in OAS1

Position of site	Flanking residues	Predicted result	SVM probability
57	EETFEV-K-FEVQKR	methylated lysine	0.576649
62	VKFEVQ-K-RQWDNP	acetyllysine	0.540308
79	FVLRSP-K-LDQEVE	methylated lysine	0.731588
100	ALGQLT-K-GYRPDS	acetyllysine	0.545806
121	ECKNLR-K-EGEFSP	methylated lysine	0.579378
138	LQRDFV-K-NRPTKL	methylated lysine	0.519411
145	NRPTKL-K-SLIRLV	methylated lysine	0.500000
164	CKEQLE-K-PLPPQY	methylated lysine	0.508369
185	AWEQGC-K-ETGFNT	methylated lysine	0.801341
185	AWEQGC-K-ETGFNT	acetyllysine	0.529851
200	GFQTVL-K-LVLKYQ	acetyllysine	0.548107
204	VLKLVL-K-YQKLCI	acetyllysine	0.505669
207	LVLKYQ-K-LCIYWE	methylated lysine	0.569388
214	LCIYWE-K-NYNSEN	methylated lysine	0.736657
233	LTKQLA-K-PRPVIL	acetyllysine	0.500000
251	TGNVAG-K-DTCSWE	methylated lysine	0.579622
273	LDYPCF-K-KRDGSP	acetyllysine	0.522832
274	DYPCFK-K-RDGSPV	methylated lysine	0.565437
274	DYPCFK-K-RDGSPV	acetyllysine	0.526673

Table 5 — Different glycosylation sites in OAS1

Position	Residue	Score	Prediction
23	NLT	0.97199888	Potential glycosylated
2	S	0.46654477	Potential glycosylated
3	S	0.29915156	Potential glycosylated
7	T	0.16234177	Potential glycosylated
8	T	0.012132963	Potential glycosylated

tool CYS_REC 37- recognizes the most probable pattern of pairs was 37-314, 47-244, 188-386.

Conserved Domain Identification

Sequence of OAS1 was submitted in BLAST for prediction of different domains in protein. It was found that it contains nucleotidyltransferase (NT) domain of 2'-5'-oligoadenylate (2-5A) synthetase (2-5OAS) and class I CCA-adding enzyme.

Prediction of Immunogenic Sites

B cell epitope prediction using IEDB analysis indicated there were 19 potential immunogenic sites in OAS1 which can act as antigen and four site act for T cell epitope given in Table 6.

Discussion

Pregnancy diagnosis is an essential component of sound reproductive management, particularly in the livestock sector where a high proportion of artificial insemination fails to produce a calf[27,28]. Early pregnancy diagnosis is crucial for decreasing the AI interval and shortening the calving interval by enabling the farmer to identify open animals so as to check and re-breed them at the earliest opportunity. Therefore, a marker molecule which can detect pregnancy at the earliest possible time is the foremost requirement to improve the reproductive efficiency in buffalo. The chemical pregnancy tests are cost effective, specifically up or down regulated, least affected due to non-animal factors, present in easily accessible body fluids and expressed over a considerable period and with no residual effect. In a chemical pregnancy test, one of the candidate molecules is OAS1. OAS1 is type I IFN with an antiviral activity which has been reported as stimulated in response to the release of another protein INF-t. Therefore, expression of OAS1 in serum level at 18 day of pregnancy can act as a suitable marker for early pregnancy detection.

The expression level of interferon stimulated genes (ISG15, OAS1, and Mx2) in buffalo (*B. bubalis*) has been reported for their potency towards ideal biomarker for detection of early pregnancy in buffalo[29-32]. The authors have reported upregulation in the expression profile of OAS1 from day 14 onwards, with the highest level at day 18 followed by gradual decline from day 21 onwards in buffalo[33]. In the present study, OAS1 mRNA amplified by semi quantitative RT-PCR, revealed specific amplification of 1170 bp product during 18 to 21 day of gestation which was indicative of increased expression level of OAS1 gene. These results were parallel to different reports which have studied the expression level of this protein and revealed a marked increase in level OAS1 protein after artificial insemination[8,11,34]. Phylogenetic analysis revealed the sequence derived has 99% similarity with predicted buffalo sequence and 94% with sheep and goat sequence. These nucleotide differences lead to change in amino acids composition which can lead to different properties in other species. OAS1 has been characterized in many species for deciphering its function as an innate immunity gene like equines, humans, murines[35,36] but no report has been there for its structural and functional prediction in buffalo. The modelled primary structure deduced by structural prediction algorithms is required to evaluate the conformational differences with other species, its cross reactivity and deciphering the activity of protein during different interaction with molecules[37]. Ramachandran plot is a way to visualize energetically allowed regions for backbone dihedral angles ψ against φ of amino acid residues in protein structure. It was found that 96.1%, residues fell in the most favored regions, 3.4% residues lies generously allowed regions and 0.5 % residues lied in the

Table 6 — Different B cell and T cell epitopic sites in OAS1

B cell epitope

NO	START	END	PEPTIDE	LENGTH
1	2	12	ELRNTPAGSLD	11
2	24	28	KFRTQ	5
3	30	30	K	1
4	57	75	VVKGGSSGKGTTLRGRSDA	19
5	89	93	EQLER	5
6	96	96	E	1
7	107	113	ACQREET	7
8	115	115	E	1
9	121	130	QKRQWDNPRA	10
10	138	145	PKLDQEVE	8
11	157	167	QLTKGYRPDSR	11
12	178	187	NLRKEGEFSP	10
13	196	203	FVKNRPTK	8
14	221	230	QLEKPLPPQY	10
15	241	254	EQGCKETGFNTAQG	14
16	274	283	KNYNSENPII	10
17	290	316	QLAKPRPVILDPADPTGNVAGKDTCSW	27
18	332	353	FKKRDGSPVGSWDVSPQEHSDL	22
19	362	383	QHYRSSPGTQFHGGASPQVEEN	22

T cell epitopes

Peptide	Length	Score
GSSGKGTTLRGRSDADLVVFLTNLTSFREQLERRGEFIEEIRRQLEACQREETFEVKFEV	60	1.4644
MELRNTPAGSLDKFIEVHLLPNEKFRTQVKEAIDIICSFLKERCFRCAPHRVRVSKVVKG	60	0.21068
RQHYRSSPGTQFHGGASPQVEENWTCTIL	29	0.18925
GSSGKGTTLRGRSDADLVVFLTNLTSFREQLERRGEFIEEIRRQLEACQREETFEVKFEV	60	1.4644

disallowed conformations in OAS1 and was the most favored structure. There was also an evaluation of percentage of alpha helices, extended strand, beta turn and random coils which revealed domination of alpha helix in OAS1 indicating folding energy for right handed is more favorable. Therefore, primary and secondary structure prediction was very useful in predicting the interaction of OAS1 with other molecules like RNAase L (innate immunity role) and IFN-t which leads to its release during viral infection as well as conceptus implantation[38]. The authors have performed homology modeling of IFN-t to reveal its interaction with other molecules involved in pregnancy detection[37]. Therefore, a structure modelled of OAS1 can be docked with other molecules already predicted for deciphering its interactions. Eukaryotic linear motifs (ELMs) are compact protein interaction sites composed of short stretches of adjacent amino acids. They are enriched in intrinsically disordered regions of the proteome and provide a wide range of functionality to proteins[39]. They play crucial roles in cell regulation and are often used by pathogens to manipulate their cellular system. This motif identification of different regions of OAS1 is helpful in identification of different parts of OAS1 protein which is involved in cell signalling processing and its interaction with other molecules.

There was a strong correlation between the stability of a protein and its physiochemical properties for predicting *in vivo* stability of a protein from its primary sequence[40]. The physiochemical properties of this protein revealed OAS1 was predicted as a very unstable and not suitable for *in vitro* heterologus expression system. The negative GRAVY value of this protein indicates that it is a protein consisting of more hydrophilic residues which may be a clue towards its secretory nature. Isoelectric point (pI) value indicated that protein will be not soluble at 8.34 pI values and buffer pH required must be far away to achieve solubility of these proteins. The theoretical pI indicates OAS1 is basic in nature which is useful for wet lab extraction (through chromatographic methods). The half-life of OAS1 was 30 hours (mammalian reticulocytes, *in vitro*), >20 hours (yeast, *in vivo*) and >10 hours (*Escherichia coli*, *in vivo*) assumed as the time taken for half of the amount of protein in a cell to disappear after its synthesis in the cell. This will be helpful in the estimation of the residual time of this protein after its expression in any system.

Many eukaryotic proteins require multiple protein translational modifications (PTMs) to reach a native, biologically active conformation. Prediction of some PTMs, such as glycosylation, methylation and disulfide bond formation, was found to significantly worsen protein amenability to soluble expression[41]. PTMs can significantly change the integral characteristics of proteins that affect their stability and solubility, such as charge, hydrophobicity, solvent accessibility, etc. Thus, in addition to the physicochemical and structural features of amino acid sequences, PTMs should also be considered as the major determinants of successful protein synthesis[42]. Phosphorylation is a vital procedure through which signaling pathways function. The removal or addition of phosphate group may result in an alteration in the function of protein and its localization[43]. Presence of 19 residues indicates different sites can be phosphorylated in OAS1 indicates its role in cell cycle, growth, cell morphology, metabolism, developmental phenomena, and virulence in OAS1[44]. Glycosylation plays a critical role in determining protein structure, folding and stability[45]. Presence of only 4 glycosylation in OAS1 protein provides suitability for expression in prokaryotic host lacking this PTM. The covalent modification of specific lysine residues in the protein, small ubiquitin-like modifiers (SUMOs) plays an essential role in the regulation of a variety of biological processes, including gene expression, DNA repair, chromosome assembly, and cellular signaling[23]. There were number of sumoylation and methylation sites in OAS1 i.e. which are very much useful in predicting structure and functional characteristics of protein.

For proper folding of any protein, disulphide bridges must be required. The misfolding in proteins is the common problem encounter during expression of any protein in heterologous system which leads to inclusion body formation[46]. The OAS1 revealed three disulpfide bridges indicating its more suitability for eukaryotic expression system than prokaryotic system such as *E. coli.* The domain identified in OAS1 was 2-5 OASs which is induced by interferon during the innate immune response to protect against RNA virus infections. In the presence of an RNA activator, 2-5 OASs catalyze the oligomerization of ATP into 2-5A. 2-5A activates endoribonuclease L, which leads to degradation of the viral RNA. 2-5 OASs are also implicated in cell growth control, differentiation, and apoptosis [47].

Immunological characterization of protein has a key role in designing of various antibody based detection kits for early pregnancy diagnosis. The present study revealed 20 immunogenic B cell epitope and 5 conformational epitopes in OAS1 which was a strong indicator of antigenic nature of these proteins. These linear and conformational 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 18-21 day after artificial insemination[48].

Conclusion

The present study concludes that OAS1 protein upregulates at 18-21 days of pregnancy in animal after artificial insemination and have desired properties for its application in pregnancy diagnosis. The prediction of various properties like structural, physiochemical, functional and immunological provides a good indication towards favorable characteristics of this molecule for being expression in any heterologous system. It was also found that OAS1 has less post translation modifications (glycosylation, phosphorylation, disulphide bridges) and number of antigenic epitopes which make it highly suitable candidate molecule for early pregnancy diagnostic kits in buffalo.

Acknowledgements

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

References

1 Drost M, Advanced reproductive technology in the water buffalo, *Theriogenology,* 68 (2007) 450-453.

2 Perera B M, BMAO: Reproductive cycles of buffalo, *Anim Reprod Sci*, 124 (2011) 194-199.

3 Thatcher W W, Meyer M D & Danet-Desnoyers G, Maternal recognition of pregnancy, *J Reprod Fertil Suppl,* 49 (1995) 15-28.

4 Spencer T E, Becker W C, George P, Mirando M A, Ogle T F *et al*, Ovine interferon-tau inhibits estrogen receptor upregulation and estrogen-induced luteolysis in cyclic ewes, *Endocrinology*, 36 (1995) 4932-4944.

5 Oliveira J F, Henkes L E, Ashley R L, Purcell S H, Smirnova N P *et al*, Expression of interferon (IFN)-stimulated genes in extrauterine tissues during early pregnancy in sheep is the consequence of endocrine IFN-tau release from the uterine vein, *Endocrinology,* 149 (2008) 1252-1259.

6 Bott R C, Ashley R L, Henkes L E, Antoniazzi A Q, Bruemmer J E *et al*, Uterine vein infusion interferon tau (IFNT) extends luteal life span in ewes, *Biol Reprod,* 82(2010) 725-735.

7 Buragohain L, Kumar R, Nanda T, Phulia S K, Mohanty A K *et al*, Serum MX2 protein as candidate biomarker for early pregnancy diagnosis in buffalo, *Reprod Domest Anim,* 51 (2016) 453-460.

8 Kizaki K, Shichijo K A, Furusawa T, Takahashi T, Hosoe M *et al*, Differential neutrophil gene expression in early bovine pregnancy, *Reprod Biol Endocrinol,* 11 (2013) 6.

9 Matsuyama S, Kojima T, Kato S & Kimura K, 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 Endocrin,* 10 (2012) 21.

10 Soumya N, Das D, Jeyakumar S, Mondal S, Mor A *et al*, Differential expression of ISG 15 mRNA in peripheral blood mononuclear cells of nulliparous and multiparous pregnant versus non-pregnant *Bos indicus* cattle, *Reprod Dom Anim,* 52 (2017) 97-106.

11 Green J C, Okamura C S, Poock S E & Lucy M C, Measurement of interferon-tau (IFN-tau) stimulated gene expression in blood leukocytes for pregnancy diagnosis within 18-20d after insemination in dairy cattle, *Anim Reprod Sci*, 121 (2010) 24-33.

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

13 Notredame C, Higgins D G & Heringa J, T-Coffee: A novel method for fast and accurate multiple sequence alignment, *J Mol Biol,* 302 (2011) 205-217.

14 Arnold K, Bordoli L, Kopp J & Schwede, T, The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling, *Bioinformatics,* 22 (2006) 195-201.

15 Luthy R, Bowie J U & Eisenberg D, Assessment of protein models with three-dimensional profiles, *Nature*, 6364 (1992) 83-85.

16 Ramachandran G N, Ramakrishnan C & Sasisekharan V, Stereochemistry of polypeptide chain configurations, *J Mol Biol*, 7 (1963) 95-99.

17 Sapay N, Guermeur Y & Deleage G, Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier, *BMC Bioinformatics,* 7 (2006) 255.

18 Dinkel H, Van Roey K, Michael S, Kumar M, Uyar B *et al*, ELM 2016-data update and new functionality of the eukaryotic linear motif resource, *Nucleic Acids Res,* 4 (2016) D294-300.

19 Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M R *et al,* Protein Identification and Analysis Tools on the ExPASy Server in: John M. Walker (Eds.), The proteomics protocols Handbook, Humana Press, (2005) 571-607.

20 Petersen B, Petersen T N, Andersen P, Nielsen M & Lundegaard C, A generic method for assignment of reliability scores applied to solvent accessibility predictions, *BMC Struct Biol,* 9 (2009) 51.

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

22 Chauhan J S, Rao A & Raghava G P S, *In silico* platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences, *Plos One*, 8 (2013) e67008.

23 Zhao Q, Yubin X, Yueyuan Z, Shuai J, Wenzhong L *et al*, GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs, *Nucleic Acids Res,* 42 (2014) W325-W330.

24 Shi S P, Qiu J D, Sun X Y, Suo S B, Huang S Y *et al*, PLMLA: prediction of lysine methylation and lysine acetylation by combining multiple features, *Mol Biosyst,* 8 (2012) 1520-1527.

25 Larsen J E, Lund O & Nielsen M, Bepipred Linear Epitope Prediction: Improved method for predicting linear B-cell epitopes, *Immunome Res,* 2 (2006) 2.

26 Rahman H A & Raghava G P S, Identification of conformational B-cell epitopes in an antigen from its primary sequence, *Immunome Res,* 6 (2010) 6.

27 Diskin M G & Sreenan J M, Fertilization and embryonic mortality rates in beef heifers after artificial insemination, *J Reprod Fertil*, 59 (1980) 463-468.

28 Oltenacu P A, Ferguson J D & Lednor A J, Economic evaluation of pregnancy diagnosis in dairy cattle: a decision analysis approach, *J Dairy Sci,* 73 (1990) 2826-2831.

29 Batra K, *In vitro* expression, molecular and immunological characterization of early pregnancy related interferon tau stimulated gene of *Bubalus bubalis,* Ph.D Dissertation, LUVAS, 2017. Hisar, Haryana, India.

30 Batra K, Nanda T, Kumar A, Kumar V, Gopal G J *et al*, Molecular cloning and expression kinetics of serum interferon stimulated gene for early pregnancy detection, *Indian J Anim Res,* 10 (2118) 18805/ijar.B-3636.

31 Batra K, Nanda T, Kumar A, Gupta A K, Kumari R, *et al*, Molecular cloning and characterization of Mx2 for early pregnancy diagnosis in *Bubalus bubalis, Proc Natl Acad Sci India Sect B Biol Sci*, (2018) https://doi.org/10.1007/s40011-018-0993-x.

32 Batra K, Kumar A, Maan S, Kumar V, Kumari R *et al*, Recombinant interferon stimulated protein 15 (rISG15) as a molecular marker for detection of early pregnancy in *Bubalus bubalis, Anim Reprod Sci,* 197 (2018) 106-116.

33 Thakur N, Singh G, Paul A, Bharati J, Rajesh G *et al*, Expression and molecular cloning of interferon stimulated genes in buffalo (*Bubalus bubalis*), *Theriogenology,* 15 (100) (2017) 50-58.

34 Gupta S S, Nanda T, Buragohain L, Kumar R & Balhara A K, OAS1 gene as potential biomarker to detect early pregnancy in buffalo, *J Cell Tissue Res,* 15 (2015) 5187-5192.

35 Rios J J, Perelygin A A, Long M T, Lear T L, Zharkikh A A *et al*, Characterization of the equine 2'-5' oligoadenylate synthetase 1 (OAS1) and ribonuclease L (RNASEL) innate immunity genes, *BMC Genomics*,7 (8) (2007) 313.

36 Eskildsen S, Justesen J, Schierup M H & Hartmann R, Characterization of the 2′-5′-oligoadenylate synthetase ubiquitin-like family, *Nucleic Acids Res*, 31 (2003) 3166-3173.

37 Joshi I, Kumar S, Kaur A, Mukhopadhyay C S & Kumar D, Homology modeling of buffalo (*Bubalus bubalis*) interferon-tau protein, *Am J Bioinform,* 1 (2012) 79-86.

38 Breda A, Valadares N F, Norberto de Souza O & Garratt R C, Protein Structure, Modelling and Applications, 2006 May 1 [Updated 2007 Sep 14]. In: Gruber A, Durham AM, Huynh C, *et al*, editors. Bioinformatics in Tropical Disease

Research: A Practical and Case-Study Approach [Internet]. Bethesda (MD): National Center for Biotechnology Information (US) 2008. Chapter A06.

39 Dinkel H, Van Roey K, Michael S, Kumar M, Uyar B *et al*, ELM-2016-data update and new functionality of the eukaryotic linear motif resource, *Nucleic Acids Res,* 44 (2016) D294-D300.

40 Guruprasad K, Reddy B & Pandit M W, Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting *in vivo* stability of a protein from its primary sequence, *Prot Eng*, 4 (1990) 155-164.

41 Tokmakov A A, Kurotani A, Takagi T, Toyama M, Shirouzu M *et al*, Multiple post-translational modifications affect heterologous protein synthesis, *J Biol Chem,* 287 (2012) 27106-27116.

42 Roy J, Posttranslational modifications and the immunogenicity of biotherapeutics, *J Immunol Res,* 15 (2016).

43 Ardito F, Giuliani M, Perrone D, Troiano G & Lo Muzio L, The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review), *Int J Mol Med,* 40 (2017) 271-280.

44 Cousin C, Derouiche A, Shi L, Pagot Y, Poncet S *et al*, Protein-serine/threonine/tyrosine kinases in bacterial signaling and regulation, *FEMS Microbiol Lett,* 346 (1) (2013) 11-19.

45 Jayaprakash N G & Surolia A, Role of glycosylation in nucleating protein folding and stability, *Biochem J,* 474 (2017) 2333-2347.

46 García-Fruitós E, Inclusion bodies: A new concept, *Microb Cell Fact,* 9 (2010).

47 Maia C J, Socorro S, Schmitt F & Santos C R, Characterization of oligoadenylate synthetase-1 expression in rat mammary gland and prostate: Effects of 17beta-estradiol on the regulation of oas1g in both tissues, *Mol Cell Biochem,* 314 (2008) 113-121.

48 Batra K & Nanda T, *In silico* characterization of interferon-stimulated gene (ISG15) as a biomarker for early pregnancy diagnosis in *Bubalus bubalis, Indian J Anim Sci,* 89 (2019) 255-263.