Nig. J. Biotech. Vol. 35 No. 2 (2018): 80-90
ISSN: 0189 1731
Available online at
http://www.ajol.info/index.php/njb/index
and www.biotechsocietynigeria.org
DOI: https://dx.doi.org/10.4314/njb.v35i2.11

Computational molecular analysis of deleterious mutations in serum amyloid A3 gene in goats and cattle

Yakubu, A.

[1]Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, P.M.B. 135, Lafia, Nasarawa State, Nigeria.

Abstract

Serum amyloid A3 (SAA3) protein found within caprine and bovine mammary epithelial cells is said to be important in disease conditions and tissue remodeling. The present investigation aimed at identifying deleterious non-synonymous single nucleotide polymorphisms (nsSNPs) in SAA3 gene of goats and cattle using an *in silico* assay. Amino acid sequence data of the protein of goats and SNPs of cattle were retrieved from the National Centre for Biotechnology Information (NCBI) database. Bioinformatics prediction tools used for the detection of deleterious nsSNPs were PROVEAN, SIFT, PolyPhe-2 and PANTHER. A total of eleven nsSNPs were obtained from the aligned sequences of goats, out of which two variants (R123G and G126D) were predicted to be deleterious by three out of the four algorithms. However, in cattle, four out of the eleven nsSNPs were found to be harmful to the transcribed protein. The two mutants in goats and R114Q in cattle were also found to decrease protein stability. Further confirmatory analysis however, revealed that variant R123G was highly deleterious as there were marked differences between it and the native protein in terms of total free energy, stabilizing residues, ordered and disordered regions of protein and secondary structure prediction. Similarly, Cmutant (a combination of R123G and G126D mutations) in goats and Dmutant (a combination of S77R, Q84K, S103W and R114Q mutations) in cattle also appeared to distort SAA3 protein structural landscape and function. The present deleterious nsSNPs when validated using wet lab experimental protocols could be important biological markers for disease detection and therapy in goats and cattle.

Keywords: protein, variant, prediction, marker, ruminants
Corresponding author's email address: abdulmojyak@gmail.com

Introduction

Mastitis, an inflammation of the mammary gland, is the most costly infectious disease of dairy ruminants worldwide (Brenaut et al., 2012) affecting milk yield and quality and the concomitant drastic reduction in farm profit. Milk contains a range of minor immune-related proteins that collectively form a significant first line of defence against pathogens, acting both within the mammary gland itself as well as in the digestive tract of the suckling neonate (Wheeler et al., 2012). Serum amyloid A (SAA) is, like C-reactive protein (CRP), an acute phase protein which, according to De Buck et al. (2016), can be utilized as a diagnostic, prognostic or therapy follow-up marker for many diseases. The mammary Serum amyloid A3 (M-SAA3) mRNA is localized to restricted caprine mammary epithelial cells (MECs). It has been reported to be expressed at a moderate level in late pregnancy; at a low level through lactation; induced early in milk stasis and expressed at high levels in most MECs of ruminants during mid to late involution including inflammation/mastitis ((Molenaar et al., 2009; Domènech et al., 2012; Domènech, 2013).

The molecular devices of the process that causes disease conditions or regulate the innate

and adaptive immune responses in livestock species may be revealed through in-depth knowledge of the structures and functions of proteins. Non-synonymous single nucleotide polymorphisms (nsSNPs) may potentially affect the function and phenotype of the encoded proteins (Samadian et al., 2017). There is a rapid growth in the identification of SNPs making it extremely difficult to evaluate the biological significance of each SNP through experimental protocols. However, first hand information could be obtained through computational, theoretical approaches to identify and investigate whether the effects of nsSNPs are deleterious to the structure and function of proteins before embarking on wet lab. experiments (Nagasundaram et al., 2015). This knowledge subsequently may be exploited in the design of better drugs to treat specific diseases (De Buck et al., 2016) especially those that are associated with the mammary glands such as mastitis.

There is dearth of information on snSNPs of SAA3 gene and the potential effects on transcribed proteins in ruminants. Therefore, the present study was embarked upon to screen deleterious amino acid substitutions of SAA3 gene in goats and cattle which may alter the conformational and functional features of the SAA3 protein.

Materials and Methods

Sequence retrieval

The amino acid sequence data on goats SAA3 gene were retrieved from the website of the National Centre for Biotechnology Information (NCBI). A total of four (4) goat sequences were obtained (Table 1). The amino acid sequence alignment of the species was carried out using ClustalW algorithm of Molecular Evolutionary Genetic Analysis (MEGA 6.0) software (Tamura et al., 2013) to obtain nonsynonymous amino acid substitutions. For cattle, apart from the amino acid sequences, fifteen single nucleotide polymorphisms (SNPs) comprising four synonymous and eleven non-synonymous amino acid substitutions were equally obtained from NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/snp) (Table 2).

Table 1: Amino acid sequence information of SAA3 gene of goats

Species	Accession Nos.	Amino Acid Nos.
Goats	ABH07648.1	131
	ABQ51184.1	131
	NP_001272566.1	128
	ABQ51197.1	128

Table 2: Single nucleotide polymorphisms record of SAA3 gene of cattle

SNP ID	Allele change	Amino acid Substitution	Substitution position	Substitution type
rs479873544	AAA ⇒ AGA	K [Lys] ⇒ R [Arg]	43	Missense
rs458097253	AGG ⇒ AAG	R [Arg] ⇒ K [Lys]	48	Missense
rs1115992367	TTC ⇒ TTT	F [Phe] ⇒ F [Phe]	54	Synonymous
rs382267063	CGT ⇒ CGC	R [Arg] ⇒ R [Arg]	57	Synonymous
rs459280223	AGT ⇒ AGG	S [Ser] ⇒ R [Arg]	77	Missense
rs443928337	AAC ⇒ GAC	N [Asn] ⇒ D [Asp]	78	Missense
rs483288056	CAG ⇒ AAG	Q [Gln] ⇒ K [Lys]	84	Missense
rs461522582	GAC ⇒ GAA	D [Asp] ⇒ E [Glu]	88	Missense
rs441703949	CCT ⇒ ACT	P [Pro] ⇒ T [Thr]	89	Missense
rs474677401	TTT ⇒ GTT	F [Phe] ⇒ V [Val]	91	Missense
rs716394599	AAG ⇒ CAG	K [Lys] ⇒ Q [Gln]	92	Missense
rs459280881	TCG ⇒ TGG	S [Ser] ⇒ W [Trp]	103	Missense
rs719093494	AAC ⇒ AAT	N [Asn] ⇒ N [Asn]	110	Synonymous
rs385589681	CGG ⇒ CAG	R [Arg] ⇒ Q [Gln]	114	Missense
rs383129737	AGC ⇒ AGT	S [Ser] ⇒ S [Ser]	115	Synonymous

Functional prediction of nonsynonymous amino acid substitutions

The functional effects of the nsSNPs of SAA3 gene in goats and cattle were predicted computationally using PROVEAN, Polyphen-2, SIFT and PANTHER. PROVEAN and Polyphen-2 models were applied as described in an earlier study (Yakubu et al., 2017) while the procedure for SIFT was highlighted in Kumar et al. (2009). SIFT uses sequence homology based algorithm to classify the effect of amino acid substitutions on protein function. PANTHER calculates the length of time (in millions of years) a given amino acid has been preserved in the lineage leading to the protein of interest (Tang and Thomas, 2016). The PANTHER thresholds used in this analysis were: "probably damaging" (time > 450my, corresponding to a false positive rate of ~0.2 as tested on HumVar), "possibly damaging" (450my > time > 200my, corresponding to a false positive rate of ~0.4) and "probably benign" (time < 200my). Where majority of the algorithms agreed on the deleterious nature of a nsSNP, further analyses were carried out to confirm or prove otherwise such a claim. Also, the deleterious nsSNPs were combined (the term 'Cmutant' was used to represent these combined deleterious nsSNPs) to exploit the effect of correlated mutations as described in Yakubu et al. (2017).

Protein stability prediction

The prediction of the effects of nsSNPs on protein stability of SAA3 gene of goats and cattle was done using I-Mutant2.0 (Capriotti et al., 2005). The free energy change (DDG) between the mutant and native proteins is predicted by I-Mutant2.0.

Tertiary structure prediction

The structural models of SAA3 of goats and cattle were constructed using the Phyre2 server (Kelly et al., 2015), which uses the alignment of hidden Markov models via HHsearch (Soding, 2005) to improve the accuracy of alignment and rate of detection. Structural similarities of alternative protein models of both species were quantified by the template used in the prediction. The models were viewed using PyMOL (DeLano, 2006) which was equally used to indicate the position of mutation on 3D structures through the Mutagenesis option.

Energy of minimization and root mean square deviation (RMSD)

Energy minimization of the modeled mutant proteins was carried out using the $3D^{refine}$ (Bhattacharya and Cheng, 2013). Its protocol refines the initial protein structures by optimizing Hydrogen Bonds network and also the atomic-level energy minimization using a combination of physics and knowledge based force fields. This force field is said to permit the evaluation of the energy of the modeled structure as well as overhaul distorted geometries through energy minimization. The mutant proteins were superimposed onto SAA3 native protein and the corresponding root mean square deviation (RMSD) values were generated using SuperPose ver 1.0 (Maiti et al., 2004).

Validation of protein structures

The proposed SAA3 protein structures of goats and cattle were validated with ERRAT and ProSA statistical softwares (Sippl, 1995; Wiederstein and Sippl, 2007). ERRAT uses characteristic atomic interaction to distinguish between regions of protein structures that are correctly and incorrectly determined (Colovos and Yeates, 1993). ProSA uses Z-scores to indicate the quality of models. It indicates that the higher the Z-score, the lower the quality of the model.

Molecular dynamic simulation

Generalized Born (GB) models, executed through The Bluues server, were used to find electrostatic differences in structures between the native and the mutant alleles of goats and cattle. The server employs the program Bluues to execute electrostatic calculations for single-atomic structures and provides options for point mutations (Walsh et al., 2012).

Identification of stabilizing residues

Stabilizing residues of the native and mutant SAA3 proteins of goats and cattle were identified with SRide. The SRide parameters were based on hydrophobicity, long-range interactions, and sequence conservation (Gromiha et al., 2004).

Prediction of ordered and disordered amino acid residues

RaptorX web server (Kallberg et al., 2012) was used to predict whether the amino acid variants are in ordered or disordered regions of the proteins. This was done for both species.

Protein-Protein Interaction

In order to predict the solvent accessibility and secondary structures in the 3D structure (Porollo and Meller, 2007) of SAA3 proteins of goats and cattle, the web-based SPPIDER Solvent accessibility based Protein-Protein Interface iDEntification and Recognition server was used.

Results

Eleven polymorphic sites were obtained from the alignment of the deduced amino acid sequences of SAA3 gene of goats (Figure 1). There was a consensus by all the algorithms, with the exception on SIFT (predicted only R123G to be harmful) in the prediction of variants R123 and G126 as being deleterious (Table 3). These two variants were therefore collectively referred to as 'Cmutant' for further confirmatory analysis. In cattle, four variants namely, S77R, Q84K, S103W and R114Q were also predicted to be deleterious by three out of the four algorithms. These four variants were thus collectively referred to as 'Dmutant'. Both variants R123G and G126D (goats) as well as R114Q (cattle) were predicted to decrease the stability of protein using the I-Mutant Suite. The respective DDG values and reliability indexes were -1.43 Kcal/mol and 8 (R123G), -1.09 Kcal/mol and 8 (G126D) as well as -1.00 Kcal/mol and 9 (R114Q).

#ABH07648.1-caprine MNLSTGIIFC FLILGVSSQG WGTFLREAGQ GAKDMWRAYR DMKEANYKGA DKYFHARGNY DAAQRGPGGA WAAKVISNAR [80]

#ABQ51197.1-caprine ....V ...E..--- .[80]

#ABQ51184.1-caprine .... [80]

#NP_001272566.1-caprine ....V ...E..---. [80]

#ABH07648.1-caprine ETIQGITDPL LKGMTRDEVR KDSKADQFAN EWGQSGKDPN HFRPAGLPDK Y [131]

#ABQ51197.1-caprine .AL..... F.....Q.. E.T..... ....G..D.... . [131]

#ABQ51184.1-caprine ....R..... ..... . [131]

#NP_001272566.1-caprine .AL..... F.....Q.. E.T..... ....G..D.... . [131]

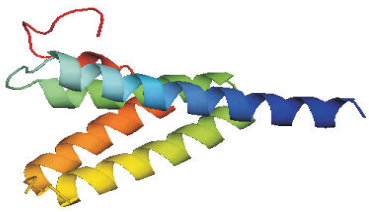
Figure 1. Aligned amino acid sequences of SAA3 protein to detect non-sysnonymous substitutions
NSites=131; Identical=. Missing=υ Indel=-

Table 3. The effect of amino acid variants on the function of SAA3 protein of goats using PROVEAN, SIFT, PolyPhen-2 and PANTHER

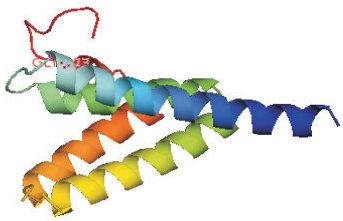
	PROVEAN		SIFT		PolyPhen-2		PANTHER	
Variant[1]	Prediction[2]	Score	Prediction[3]	Score	Prediction[4]	Score	Prediction[5]	Time
Goats								
A70V	N	-0.680	T	0.70	B	0.001	P	455
K74E	N	-1.253	T	0.44	B	0.001	S	220
T82A	N	-0.415	T	0.50	B	0.000	B	2
I83L	N	-0.104	T	1.00	B	0.000	B	2
L91F	N	0.745	T	0.69	B	0.000	N	-
E98Q	N	-0.083	T	0.36	B	0.000	N	-
K101E	N	2.852	T	1.00	B	0.000	N	-
S103T	N	-1.036	T	0.43	B	0.005	P	455
Q114R	N	3.249	T	1.00	B	0.000	N	-
R123G	D	-5.960	D	0.00	P	0.630	P	910
G126D	D	-5.139	T	0.05	P	0.630	P	910
Cattle								
K43R	N	0.941	T	1.00	B	0.002	B	2
R48K	N	1.669	T	1.00	B	0.000	B	2
S77R	D	-4.711	D	0.00	S	0.944	P	910
N78D	N	0.833	T	1.00	B	0.001	B	2
Q84K	D	-3.296	D	0.02	S	0.581	P	910
D88E	N	-0.813	D	0.02	B	0.101	B	176
P89T	N	0.065	T	0.50	B	0.071	B	85
F91V	N	-0.979	T	0.32	B	0.003	B	85
K92Q	N	-0.137	T	0.38	B	0.144	B	85
S103W	D	-4.781	D	0.00	P	0.998	P	455
R114Q	D	-3.249	T	0.05	S	0.884	P	910

1 A = alanine, V = valine, K = lysine, E = glutamic acid, T = threonine, I = isoleucine, L = leucine, F = phenylalanine, Q = glutamine, S = serine, R = arginine, G = glycine, D = aspartic acid; N = asparagine; P = proline; W = tryptophan
2 N = neutral; D = deleterious
3 T = tolerated; D = deleterious
4 B = benign; P = probably damaging, S = possibly damaging
5 P = probably damaging; S = possibly damaging; B = probably benign; N = not scored

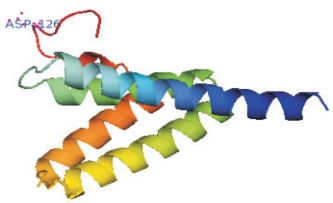
105 residues (80% of SAA3 amino acid sequence) with 100.0% confidence level, has been modelled by the single highest scoring template for the native protein, R123G and G126D variants including the Cmutant in goats (Figure 2). However, in cattle, 104 residues (73% of sequence) have been modelled with 100.0% confidence level for the native protein including variants S77R, Q84K, S103W and R114Q, respectively. These percentage values are indications of the correct prediction of the SAA3 3D structures of the native and mutants in both species.



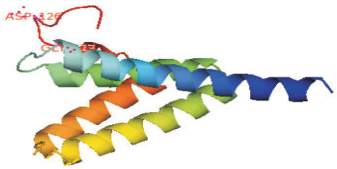
a



b

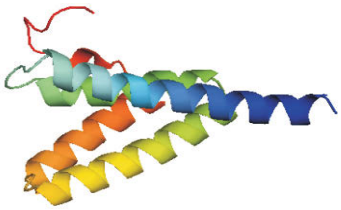


c

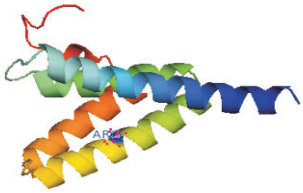


d

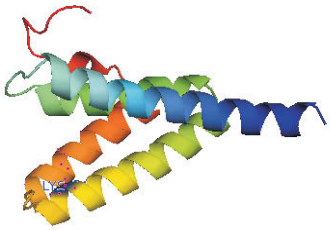
Figure 2. The predicted 3D structures of the native and mutant SAA3 proteins of goats
a: native SAA3 protein
b: SAA3 protein due to mutant R123G [Arginine (Arg) to Glycine (Gly) at position 123]
c: SAA3 protein due to mutant G126D [Glycine (Gly) to Aspartic Acid (Asp) at position 126]
d: SAA3 protein due to Cmutant (Combination of R123G and G126D)



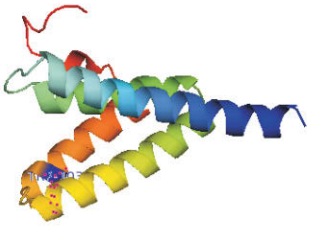
e



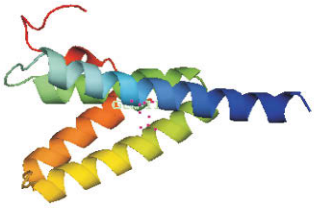
f



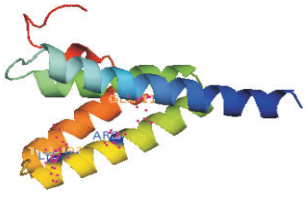
g



h



l



j

Figure 3. The predicted 3D structures of the native and mutant SAA3 proteins of cattle
e: native SAA3 protein of cattle
f: SAA3 protein due to mutant S77R [Serine (Ser) to Arginine (Arg) at position 77]
g: SAA3 protein due to mutant Q84K [Glutamine (Gln) to Lysine (Lys) at position 84]
h: SAA3 protein due to mutant S103W [Serine (Ser) to Tryptophan (Trp) at position 103]
i: SAA3 protein due to mutant R114Q [Arginine (Arg) to Glutamine (Gln) at position 114]
j: SAA3 protein due to Dmutant (Combination of S77R, Q84K, S103W and R114Q)

The values of RMSD of the mutants of both caprine and bovine species obtained from SuperPose are shown in Table (4). RMSD value of 0.19 for variants R123G and Cmutant though less than the threshold of > or = 2.0 for a variant to effect structural change, is higher than the value of 0.05 recorded for the native and G126D. In cattle, only substitution R114Q recorded RMSD value of less than 0.1

Table 4: Root mean square deviation values of the native and mutant proteins

Variant	Root mean square deviation (RMSD)			
	Alpha Carbon	Backbone	Heavy	All
Goats				
R123G	0.15	0.15	0.19	0.19
G126D	0.05	0.04	0.05	0.05
Cmutant	0.15	0.15	0.19	0.19
Cattle				
S77R	0.09	0.09	0.13	0.13
Q84K	0.05	0.05	0.13	0.13
S103W	0.08	0.08	0.12	0.12
R114Q	0.06	0.06	0.09	0.09
Dmutant	0.11	0.11	0.18	0.18

In goats, the native protein and the amino acid substitutions R123G, G126D and Cmutant all recorded 97.94 overall quality factor as revealed by ERRAT. Overall quality factor for the native protein and variants Q84R and S103W in cattle was 98.96 whereas values of 90.63, 97.92 and 100.00 were recorded for variants S77R, R114Q and Dmutant, respectively. However, the ProSA Z-scores of R123G and Cmutant (-2.57 and -2.58) were less negative than that of the native and G126D proteins (-2.72 in each case). Similarly, variants R114Q (-2.63) and Dmutant (-2.71) have less negative ProSA Z-scores compared to S77R (-2.83), Q84K (-2.79), S103W (-2.79) and the native protein (-2.78).

The molecular dynamic simulation revealed that the native proteins and the variants in both species differed in born self energy, coulomb energy, electrostatic solvation energy and total energy (Table 5).

Table 5: Energy differences between the native protein and mutants

Substitution	Born self energy (kJ)	Coulomb energy (kJ)	Electrostatic solvation energy (kJ/mol)	Total energy (kJ/mol)
Goats				
Native	-4799.23	-30084.77	-654.54	-29923.33
R123G	-4726.46	-29212.49	-757.66	-29149.12
G126D	-4863.00	-30196.36	-766.26	-30141.67
Cmutant	-4790.80	-29288.99	-903.29	-29366.27
Cattle				
Native	-4949.43	-31233.41	-512.46	-30925.49
S77R	-5003.17	-31886.52	-402.79	-31472.39
Q84K	-4934.56	-31113.68	-454.61	-30748.75
S103W	-4936.20	-31202.60	-548.12	-30933.92
R114Q	-4900.90	-30794.98	-491.40	-30465.87
Dmutant	-4925.79	-31273.76	-421.94	-30882.85

There was a single stabilizing amino acid residue (Alanine) at position 50 for the SAA3 native protein and mutant G126D (Table 6). However, no stabilizing residue was obtained in case of the substitutions R12G and Cmutant. In cattle, the native protein and all the mutants have Alanine at position 50.

Table 6: Variation in the stabilizing residues of the native and mutant proteins of SAA3 gene

Substitution	Residue	Conservation score	Hydrophobicity	Long-range order
Goats				
Native	ALA_50	6	22.58	0.03810
R123G	–	–	–	–
G126D	ALA_50	6	22.58	0.03810
Cmutant	–	–	–	–
Cattle				
Native	ALA_50	5	21.79	0.03846
S77R	ALA_50	5	21.79	0.03846
Q84K	ALA_50	5	21.79	0.03846
S103W	ALA_50	6	21.79	0.03846
R114Q	ALA_50	5	21.79	0.03846
Dmutant	ALA_50	5	21.79	0.03846

In goats, the mutant R123G was found to be a change from ordered (R) to disordered (G) amino acid residue. However, the mutant G126D was a change from disordered (G) to ordered (D) residue. In cattle, all the mutants (S77R, Q84K, S103W and R114Q) were found as ordered residues. All predictions were done at a false positive rate threshold of 5% in both species.

The native and the mutants varied in the number of H-Helix and E-Beta Strand. While substitution R123G has a loss of residue in H-Helix, it however, has a gain of residue in E-Beta Strand compared to the native protein (Table 7). The Cmutant H-Helix was higher by a single residue, but in terms of E-Beta Strand was lower by a single residue. The native and the mutants all have equal number of C-Coil residues. In cattle, variant Q84K and Dmutant were higher by two residues in C-Coil, one residue in E-Beta Strand and lower by three residues in H-Helix compared with the native protein. Apart from variant S77R, every other mutant differed from the native protein in terms of secondary structure configuration. Substitutions R123G and G126D (goats), S77R and S103W (cattle) were found as non-interfacial residues in soluble domain. However, Q84K and R114Q were exposed (interfacial residues in soluble domain).

Table 7: Secondary Structure prediction using SPPIDER software

Variants	Secondary Structure		
	H-Helix	E-Beta Strand	C-Coil
Goats			
Native	55	2	48
R123G	54	3	48
G126D	55	2	48
Cmutant	56	1	48
Cattle			
Native	57	4	43
S77R	57	4	43
Q84K	54	5	45
S103W	56	5	43
R114Q	56	5	43
Dmutant	54	5	45

Discussion

It has been reported that in non-human mammals, SAA3 is the main SAA form being expressed extrahepatically (Upragarin et al., 2005) with evidence in the colostrum of ruminant animals (McDonald et al., 2001). This gene may play a role in the protection of the mammary gland during remodelling and infection (Molenaar et al., 2009). The high number of polymorphic sites predicted to be neutral at SAA3 locus of cattle and goats in the present study could be an indication of high conservatism. This is in line with the submission of Domènech Guitart (2013) that the protein sequences of SAA are highly conserved with a wide range of mammals exhibiting high homology. This conservation has been maintained through the evolution of eutherian mammals (Uhlar and Whitehead, 1999), indicating that the SAA3 locus may likely have biological functions of importance. The consensus in the prediction of R123G and G126D (goats) including S77R, Q84K, S103W and R114Q (cattle) as being deleterious may be a pointer to their pathological phenotypic consequences. The high reliability indexes obtained in the present study when the substitutions R123G and G126D including R114Q were subjected to stability test further indicate that the three variants could be disease-related mutations.

Variants R123G and Cmutant (goats) as well as Dmutant (cattle) appear to have more propensity to alter protein structural landscape than other mutants though below RMSD value threshold. Based on the fact that they have less negatve Z-scores, it seems variants R123G and Cmutant including R114Q and Dmutant (-2.71) are more disposed to affecting the protein structure. This is consistent with the report of Saha et al. (2013) that the model quality is better the more negative the Z-score is.

There is every tendency that R123G, Q84K, R114Q, Cmutant and Dmutant changes could affect protein conformation and biological roles by virtue of their less negative total free energy. In a related study, Alberts et al. (2002) reported that the native structure or conformation of a protein generally is enhanced by free energy minimization. However, the denatured state of proteins, as observed in the mutants of the present study is characterized by high conformational entropy, which could be attributed to loss of interactions by the natives. This may destabilize the protein molecules and their subsequent functions.

The identification of stabilizing residues could be used as potential candidates for studying protein folding and stability. This is because a single incorrectly predicted mutation, apart from jeopardizing the stability of the entire protein, can counteract a larger number of stabilizing mutations (Broom et al., 2017). The conversion of R123G from ordered to disordered residue in the mutant may have structural and functional effect on the SAA3 protein. Arginine (R) is a basic-polar amino acid while glycine (G) is non-polar. This change from a charged amino acid/hydrophilic to non-polar/neutral state could disturb the ionic interaction in the native protein, thereby affecting the structural configuration. This is congruous to the submission of Yakubu et al. (2017). Protein fragments that are not well ordered in the crystal have been reported to simply not visible in electron density, as a result of which they are not built into the final model (Zhang et al., 2007). Disordered regions of protein could have specific structural and amino acid affinities (Kumari et al., 2015) compared to the ordered regions. Characterization of the physico-chemical properties of protein-protein interactions is very useful because key biological processes such as antigen-antibody recognition, hormone-receptor binding and signal transduction are said to be regulated through proteins' association and dissociation (Zen et al., 2010).

The varying configurations of the native and the SAA3 protein variants R123G, Q84K, Cmutant and Dmutant could disturb protein folding and hence its structure and function. This is because the 20 types of amino acids in the three distinct secondary structures (helix, beta strand, and loop) have been reported to provide important information on the interaction preferences of amino acids in the folding of proteins (Jha et al., 2010). The non-interfacial disposition of the mutant R123G, S77R and S103W could also affect their level of interaction which could possibly exert a disruptive effect on dimerization. According to Zhan and Lazaridis (2009), non-interfacial ionizable residues can influence dimer association by inducing large conformational changes.

Conclusion

The present findings provide information on the molecular basis of the potential deleterious effects of R123 and Cmutant on protein structures and functions of SAA3 gene of goats. In cattle, the effect of Dmutant on protein configuration was more pronounced. When these variants are experimentally confirmed in future web lab and pathogenic population-based association studies, may be exploited in unraveling the genetic resistance/susceptibility to mammary diseases in goats and cattle.

Acknowledgement

Facilities for this research were provided by the Department of Animal Science, Faculty of Agriculture, Nasarawa State, University, Keffi.

References

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002.

Bhattacharya, D. and Cheng, J. (2013). 3Drefine: Consistent Protein Structure Refinement by Optimizing Hydrogen Bonding Network and Atomic-Level Energy Minimization. Proteins, 81(1):119–31.

Brenaut, P., Bangera, R., Bevilacqua, C., Rebours, E., Cebo, C. and Martin, P. (2012). Validation of RNA isolated from milk fat globules to profile mammary epithelial cell expression during lactation and transcriptional response to a bacterial infection. J. Dairy Sci. 95: 6130–6144.

Broom, A., Jacob, J., Trainor, K. and Meiering, E.M. (2017). Computational tools help improve protein stability but with a solubility tradeoff. The J. Biol. Chem. 292: 14349-14361.

Capriotti, E., Fariselli, P. and Casadio, R. (2005). I-Mutant2.0: Predicting stability changes upon mutation from the protein sequence or structure. Nucl. Acids Res. 33: 306-310.

Colovos, C. and Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. Protein Sci. 2 (9): 1511-1519.

De Buck, M., Gouwy, M., Wang, J. M., Van Snick, J., Opdenakker, G., Struyf, S. and Van Damme, J. (2016). Structure and expression of different serum amyloid A (SAA) variants and their concentration-dependent functions during host insults. Curr. Med. Chem. 23 (17): 1725–1755.

DeLano, W. L. (2006). PyMOL Executable Build (v0.99); DeLano Scientific: South San Fransisco, CA 94080, USA.

Domènech, A., Raynes, J. G., Rodríguez, E. G., Arís, A., Bach, A. and Serrano, A. (2012). Recombinant expression of goat milk serum amyloid A: preliminary studies of the protein and derived peptides on macrophage phagocytosis. Protein Pept. Lett. 19 (3): 299-307.

Domènech Guitart, A. (2013). Analysis of the functional roles of mammary serum amyloidA3 protein. A Ph.D thesis submitted to the Department de Genètica i Microbiologia Facultat de Biociències Universitat Autònoma de Barcelona, June 2013. 135 pp.

Gromiha, M. M., Pujadas, G., Magyar, C., Selvaraj, S. and Simon, I. (2004). Locating the stabilizing residues in (alpha/beta)8 barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation. Proteins 55: 316-329.

Jha, A.N., Vishveshwara, S. and Banavar, J.R. (2010). Amino acid interaction preferences in proteins. Protein Sci. 19 (3): 603–616.

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H. and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. Nature Protocols 7: 1511–1522.

Kelley, L. A., Mezulis, S., Yates, C. M.,Wass, M. N. and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10: 845-858.

Kumar, P., Henikoff, S. and Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm, Nat. Protoc. 4: 1073–1081.

Kumari, B., Kumar, R. and Kumar, M. (2015). Low complexity and disordered regions of proteins have different structural and amino acid preferences. Mol. Biosyst. 11(2): 585-594.

McDonald, T. L., Larson, M. A., Mack, D. R. and Weber, A. (2001). Elevated extrahepatic expression and secretion of mammary-associated serum amyloid A 3 (M-SAA3) into colostrums. Vet. Immunol. Immunopathol. 83 (3-4): 203-211.

Maiti, R., Van Domselaar, G. H., Zhang, H. and Wishart, D. S. (2004) "SuperPose: a simple server for sophisticated structural superposition" Nucleic Acids Res. 32 (Web Server issue): W590W594.

Molenaar, A. J., Harris, D. P., Rajan, G. H., Pearson, M. L., Callaghan, M. R., Sommer, L., Farr V. C., Oden, K. E., Miles, M. C., Petrova, R. S., Good, L. L., Singh, K., McLaren, R. D., Prosser, C. G., Kim, K. S., Wieliczko, R. J., Dines, M. H., Johannessen, K. M., Grigor, M. R., Davis, S. R. and Stelwagen, K. (2009). The acute-phase protein serum amyloid A3 is expressed in the bovine mammary gland and plays a role in host defence. Biomarkers 14 (1): 26-37.

Nagasundaram, N., Zhu, H., Liu, J., Karthick, V., Doss, C.G.P., Chakraborty, C. and Chen, L. (2015). Analysing the effect of mutation on protein function and discovering potential inhibitors of CDK4: Molecular modelling and dynamics studies. PLoS ONE 10 (8): e0133969.

Porollo, A. and Meller, J. (2007). Prediction-based fingerprints of protein-protein interactions. Prot. Struct. Funct. Bioinform. 66: 630–645.

Saha, C., Polash, A. H., Islam, M. T. and Shafrin, F. (2013). In silico prediction of structure and functions for some proteins of male-specific region of the human Y chromosome. Interdiscip. Sci. Comput. Life Sci. 5: 258-269.

Samadian, E., Gharaei, R., Colagar, A. H. and Sohrabi, H. (2017). Computational study of putative functional variants in human kisspeptin. J. Genet. Eng. Biotech. 15 (2): 419-422.

Sippl, M. J. (1995). Knowledge-based potentials for proteins. Curr. Opin. Struct. Biol. 5: 229-235.

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. Bioinformatics, 21: 951-960.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. 30: 2725-2729.

Tang, H. and Thomas P. D. (2016). PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. Bioinformatics 32 (14): 2230-2232.

Uhlar, C. M. and Whitehead, A. S. (1999). Serum amyloid A, the major vertebrate acute-phase reactant. The FEBS J. 265: 501-523.

Upragarin, N., Landman, W. J., Gaastra, W. and Gruys, E. (2005). Extrahepatic production of acute phase serum amyloid A. Histol. Histopathol. 20 (4): 1295-1307.

Walsh, I., Minervini, G., Corazza, A., Esposito, G., Tosatto, S. C. E. and Fogolari, F. (2012). Bluues Server: electrostatic properties of wild-type and mutated protein structures. Bioinformatics 28 (16): 2189-2190.

Wheeler, T. T., Smolenski, G. A., Harris, D. P., Gupta, S. K., Haigh, B. J., Broadhurst, M. K., Molenaar, A. J. and Stelwagen, K. (2012). Host-defence-related proteins in cows' milk. Animal 6 (3): 415-422.

Wiederstein, M. and Sippl. M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35(suppl 2): W407-W410.

Yakubu, A., De Donato, M. and Imumorin, I. G. (2017). Modelling functional and structural impact of non-synonymous single nucleotide polymorphisms of the DQA1 gene of three Nigerian goat breeds. S. Afr. J. Anim. Sci. 47 (2): 146-156.

Zen, A., Micheletti, C., Keskin, O. and Nussinov, R. (2010). Comparing interfacial dynamics in protein-protein complexes: an elastic network approach. BMC Struct. Biol. **10**: 26.

Zhang, J. and Lazaridis, T. (2009). Transmembrane helix association affinity can be modulated by flanking and noninterfacial residues. Biophys. J. 96: 4418-4427.

Zhang, Y., Stec, B. and Godzik, A. (2007). Between order and disorder in protein structures analysis of "dual personality" fragments in proteins. –Structure 15(9): 11411147.