ABG -15

In-Silico Analysis of Deleterious Mutations in Serum Amyloid A3 Gene in Goats

A. Yakubu
Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, PMB 135 Lafia, Nasarawa State, Nigeria.
Corresponding authors: A. Yakubu; E-mail: abdulmojyak@gmail.com

Abstract
Serum amyloid A3 (SAA3) protein found within caprine 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 using an *in silico* assay. Amino acid sequence data of the protein were retrieved from the National Centre for Biotechnology Information (NCBI) website and aligned to obtain amino acid substitutions. Bioinformatics prediction tools used for the detection of deleterious nsSNPs were PROVEAN, SIFT, PolyPhe-2 and PANTHER. A total of eleven nsSNPs were obtained, out of which two variants (R123G and G126D) were predicted to be deleterious by three out of the four algorithms. The mutants 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) was also found 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.

Keywords: protein, variant, prediction, marker, goats

Introduction

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). 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).

Firsthand 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. This knowledge may be exploited in the design of better drugs to treat specific diseases especially those that are associated with the mammary glands. 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 which may alter the conformational and functional features of the SAA3 protein.

Materials and Methods

A total of four (4) amino acid sequence data on goats SAA3 gene were retrieved from the website of the National Centre for Biotechnology Information (NCBI). The amino acid sequence alignment of the species was carried out using ClustalW algorithm of MEGA 6.0 software. The functional effects of the non synonymous amino acid substitutions (nsSNPs) of SAA3 gene in goats were predicted computationally using PROVEAN, Polyphen-2, SIFT and PANTHER. Also, the deleterious nsSNPs were combined (the term 'Cmutant' was used to represent these combined deleterious nsSNPs) to exploit the effect of correlated mutations (Yakubu *et al*., 2017). The prediction of the effects of nsSNPs on protein stability of SAA3 gene of goats was done using I-Mutant2.0. Energy minimization of the modeled mutant proteins was carried out using the $3D^{refine}$. 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. 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. Stabilizing residues of the native and mutant SAA3 proteins of goats were identified with SRide.

RaptorX web server was used to predict whether the amino acid variants are in ordered or disordered regions of the proteins. In order to predict the solvent accessibility and secondary structures in the 3D structure of SAA3 proteins of goats, the web-based SPPIDER was used.

Results and Discussion

Eleven polymorphic sites were obtained from the alignment of the deduced amino acid sequences of SAA3 gene of goats. 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 1). These two variants were therefore collectively referred to as 'Cmutant' for further confirmatory analysis. Both variants R123G and G126D were predicted to decrease the stability of protein using the I-Mutant Suite. The SAA3 3D structures of the native and mutants were correctly predicted (80% of SAA3 amino acid sequence). RSMD value of 0.19 for variants R123G and Cmutant is close to the threshold of ≥ 2.0 for a variant to effect structural change. However, the value of 0.05 recorded for the native and G126D mutant was far less than the threshold. The molecular dynamic simulation revealed that the native protein and the variants R123G, D126D and Cmutant differed in born self energy, coulomb energy, electrostatic solvation energy and total energy (Table 2). There was a single stabilizing amino acid residue (Alanine) at position 50 for the SAA3 native protein and mutant G126D. However, no stabilizing residue was obtained in case of the substitutions R12G and Cmutant. 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. The native and the mutants varied in the number of H-Helix and E-Beta Strand. Substitutions R123G and G126D were found as non-interfacial residues in soluble domain.

Table 1: The effect of amino acid variants on the function ofSAA3 protein of goats using PROVEAN, SIFT, PolyPhen-2 and PANTHER

	PROVEN		SIFT		PolyPhen-2		PANTHER	
Variant[1]	Prediction[2]	Score	Prediction[3]	Score	Prediction[4]	Score	Prediction[5]	Time
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

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;
2 N = neutral; D = deleterious
3 T = tolerated; D = deleterious
4 B = benign; P = possibly damaging
5 P = probably damaging; S = possibly damaging; B = probably benign; N = not scored

Table 2: 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)
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

The few polymorphic sites obtained at SAA3 locus of goats in the present study is an indication of high conservatism. This is in line with the submission of Domènech Guitart (2013). The consensus in the prediction of R123G and G126D as being deleterious may be a pointer to their pathological phenotypic consequences. There is every tendency that R123G and Cmutant 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. The varying configurations of the native and the SAA3 protein variants R123G and Cmutant could disturb protein folding and hence its structure and function.

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. When this is 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.

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.

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.

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.

Domènech Guitart, A. (2013). Analysis of the functional roles of Mammary Serum AmyloidA3 protein. Ph.D Thesis, Department de Genètica i Microbiologia Facultat de Biociències Universitat Autònoma de Barcelona, June 2013. 135 pp.

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 defense. *Biomarkers,* 14(1): 26-37.

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.

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.