# Characterizing the polymorphism K232A of the diacylglycerol-acyltransferase-1 lipogenic enzyme of bovine *Bos taurus* using *in silico* comparative protein prediction analyses

*F. F. P. Costa[a], J. F. Nogueira[a], J. J. S. Gouveia[a] and J. E. Gabriel[a]\**

[a]Universidade Federal do Vale do São Francisco – UNIVASF, Campus de Ciências Agrárias, Rodovia BR 407, Km 12, Projeto de Irrigação Nilo Coelho, CEP 56300-000, Petrolina, PE, Brazil

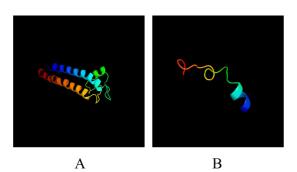*e-mail: jane.gabriel@univasf.edu.br

The diacylglycerol-acyltransferase-1 lipogenic enzyme (DGAT1) is a multi-pass transmembrane protein that catalyzes the conversion of diacylglycerol and fatty acyl CoA to triacylglycerol or retinol (LIU et al., 2012). The lysine to alanine polymorphism in the DGAT1 (K232A) is directly associated with 50% of the genetic variation in milk-fat percentage and exerts a strong effect on milk fatty acid composition (Schennink et al., 2008). Within this perspective, this study aimed to predict potential structural and functional effects of the polymorphism K232A on DGAT1 protein sequences of bovine *Bos taurus* using *in silico* comparative modelling prediction analyses.

Two sequences of the DGAT1 protein of bovine *Bos taurus* were analyzed in this study: the non-polymorphic DGAT1 sequence (ID:Q8MK44) searched and selected from a non-redundant protein sequence database UniProtKB/Swiss-Prot, and polymorphic DGAT1 sequence harboring the mutation K232A. Functional predictions of these proteins were evaluated using SIFT and PolyPhen-2 tools that infer the effect of an amino acid substitution on the structure and function of a protein using sequence homology (Kumar et al., 2009; Adzhubei et al., 2010). The prediction of differences in the protein stability due to mutation was also estimated from Gibbs free energy change (DDG) using I-Mutant version 2.0 (Capriotti et al., 2005). *In silico* comparative modelling prediction analyses of the tertiary and quaternary structures of the non-polymorphic and polymorphic DGAT-1 proteins were predicted using Phyre2 (Kelley et al., 2015) and Swiss-Model (Biasini et al., 2014) servers that build 3D theoretical models for protein fold recognition based on evolutionary related proteins and homology modelling. All computational biology tools selected in current study have been largely described in the literature as sufficient and reliable evaluative tools for *in silico* characterization of multiple target proteins.

The SIFT and PolyPhen-2 results demonstrated that the mutation K232A was predicted to be tolerated with a score of 0.19 and benign with a score to 0.001, 0.99 sensitivity and 0.15 specificity. Furthermore, DDG value resulting of the polymorphism K232A was 0.78, indicating an increase in the structural stability of the polymorphic DGAT1. In Phyre2 analyses the modelling prediction for

the non-polymorphic DGAT1 sequence targeted directly to Chain A of the membrane channel protein WSK3, a water-soluble analogue of the potassium channel KcsA, modelling 96 residues (20% of the sequence) with 23.9% confidence (Figure 1A). Values of the global modelling quality estimation (GMQE and QMEAN) of the target-template alignment were 0.06 and -3.57, respectively. Notably, the modelling prediction for the polymorphic DGAT1 sequence matched to Chain A of a stromal interaction molecule 2, modelling 21 residues (4% of the sequence) with 23.0% confidence (Figure 1B), and 0.04 GMQE and -6.27 QMEAN values. According to Kelley et al. (2015), 3D structure protein modellings may be very useful as long as the confidence is high, even at very low sequence identities (<15%). Both non-polymorphic and polymorphic DGAT1 proteins were topologically predicted as transmembrane helices (data not shown). Although QMEAN value has indicated very low Z-scores in these *in silico* analyses, the quality estimates for membrane proteins, such as DGAT1 protein, differ considerably from those of soluble proteins due to their physico-chemical properties.

Regarding to post-translational protein modifications, Swiss-Model analyses revealed the direct binding of the N-acetyl-D-glucosamine (NAG) and 1,2-diacyl-SN-glycero-3-phosphocholine (PC1) ligands to non-polymorphic DGAT1 protein target (data not shown). Nevertheless, none binding of ligands to protein target was revealed in the predicted models for the polymorphic DGAT1 protein. The NAG is an ubiquitous post-translational protein modification, consisting of a single N-acetylglucosamine moiety linked to serine or threonine residues on nuclear and cytoplasmic proteins that is involved in sensing the nutrient status of the surrounding cellular environment (Graham et al., 2011). Alternatively, a phosphatidylcholine PC1 is a glycerol phosphatide in which the hydroxy group of choline is esterified with the phosphate group of phosphatidic acid, exerting direct functional roles on the phospholipid metabolism (Zhu et al., 2001).

The findings described in current study demonstrate that the polymorphism K232A provokes an increase in the structural stability of the polymorphic protein and drastic alterations of the 3D structure protein homology

*Braz. J. Biol. 2018, vol. 78, no. 2, pp. 389-390*

389

**Figure 1.** *In silico* comparative protein prediction analyses of the diacylglycerol-acyltransferase-1 lipogenic enzyme of bovine *Bos taurus* using Phyre2 protein fold recognition server. (A) Non-polymorphic DGAT1 sequence, model dimensions (Å): **X**:29.649, **Y**:54.608, **Z**:43.339; (B) Polymorphic DGAT1 sequence harboring the polymorphism K232A, model dimensions (Å): X:36.965, Y:15.920, Z:15.699. Rainbow-color-code indicates the protein chain from blue at the N-terminus to red at the C-terminus. Dimensions of the three-dimensional modellings were computed along the specified projection axis X:Y:Z for this algorithm in particular.

modelling compared to non-polymorphic DGAT1 protein (Figure 1). Thus, efforts should be focused on development of comparative protein structure modelling algorithms usable to minimize typical limitations associated with remarkable complexity and variability of transmembrane proteins, such as the number of torsion angles in disallowed regions and exposed hydrophobic residues, hydrogen bonds and exposed polar or charged residues (Zhu et al., 2001). Members of a same family of proteins containing non-synonymous replacements of amino acids have frequently predicted structural and functional differences by computational biology analyses (Gabriel et al., 2013). Substitutions, deletions or insertions of amino acid residues are extensively typical in the outer surface of the proteins, particularly in the loops, structurally variable regions most exposed to the protein. Within this context, the presence of the non-synonymous mutation K232A in the DGAT1 protein seems to be sufficient to promote an expressive alteration in its theoretical modelling pattern for folding protein in comparison to non-polymorphic polypeptide (Figure 1).

In conclusion, such findings describing the first reports in the literature of *in silico* analyses for folding DGAT1 protein, indicate that the polymorphism K232A exerts a benign and tolerant effect on the functional activity and structural stability of the DGAT1 protein, although the protein modelling prediction has been sensibly affected due to mutation. Furthermore, the predicted binding of ligands associated potentially with lipid metabolism to non-polymorphic DGAT1 protein target may suggest its susceptibility to share posttranslational modifications that affect the metabolic pathways modulating the production and composition of the milk in bovine.

## References

ADZHUBEI, I.A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V.E., GERASIMOVA, A., BORK, P., KONDRASHOV, A.S. and SUNYAEV, S.R., 2010. A method and server for predicting damaging missense mutations. *Nature Methods*, vol. 7, no. 4, pp. 248-249. PMid:20354512. http://dx.doi.org/10.1038/nmeth0410-248.

BIASINI, M., BIENERT, S., WATERHOUSE, A., ARNOLD, K., STUDER, G., SCHMIDT, T., KIEFER, F., CASSARINO, T.G., BERTONI, M., BORDOLI, L. and SCHWEDE, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, vol. 42, no. W1, pp. 195-201. PMid:24782522. http://dx.doi.org/10.1093/nar/gku340.

CAPRIOTTI, E., FARISELLI, P. and CASADIO, R., 2005. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research*, vol. 33, pp. W306-W310. PMid:15980478. http://dx.doi.org/10.1093/nar/gki375.

GABRIEL, J.E., FIGUEIREDO, D.D. and FARIAS, R.P., 2013. Revealing highly conserved regions in the E6 protein among distinct human papillomavirus types using comparative analysis of multiple sequence alignments. *Brazilian Journal of Biology = Revista Brasileira de Biologia*, vol. 73, no. 2, pp. 449-450. PMid:23917577. http://dx.doi.org/10.1590/S1519-69842013000200030.

GRAHAM, M.E., THAYSEN-ANDERSEN, M., BACHE, N., CRAFT, G.E., LARSEN, M.R., PACKER, N.H. and ROBINSON, P.J., 2011. A novel post-translational modification in nerve terminals: O-linked N-acetylglucosamine phosphorylation. *Journal of Proteome Research*, vol. 10, no. 6, pp. 2725-2733. PMid:21500857. http://dx.doi.org/10.1021/pr1011153.

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*, vol. 10, no. 6, pp. 845-858. PMid:25950237. http://dx.doi.org/10.1038/nprot.2015.053.

KUMAR, P., HENIKOFF, S. and NG, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, vol. 4, no. 8, pp. 1073-1081. PMid:19561590. http://dx.doi.org/10.1038/nprot.2009.86.

LIU, Q., SILOTO, R.M., LEHNER, R., STONE, S.J. and WESELAKE, R.J., 2012. Acyl-CoA:diacylglycerol acyltransferase: molecular biology, biochemistry and biotechnology. *Progress in Lipid Research*, vol. 51, no. 4, pp. 350-377. PMid:22705711. http://dx.doi.org/10.1016/j.plipres.2012.06.001.

SCHENNINK, A., HECK, J.M.L., BOVENHUIS, H., VISKER, M., VAN VALENBERG, H.J. and VAN ARENDONK, J.A., 2008. Milk fatty acid unsaturation: genetic parameters and effects of stearoyl-CoA desaturase (SCD1) and acyl CoA: diacylglycerol acyltransferase 1 (DGAT1). *Journal of Dairy Science*, vol. 91, no. 5, pp. 2135-2143. PMid:18420645. http://dx.doi.org/10.3168/jds.2007-0825.

ZHU, H., BILGIN, M., BANGHAM, R., HALL, D., CASAMAYOR, A., BERTONE, P., LAN, N., JANSEN, R., BIDLINGMAIER, S., HOUFEK, T., MITCHELL, T., MILLER, P., DEAN, R.A., GERSTEIN, M. and SNYDER, M., 2001. Global analysis of protein activities using proteome chips. *Science*, vol. 293, no. 5537, pp. 2101-2105. PMid:11474067. http://dx.doi.org/10.1126/science.1062191.

390

*Braz. J. Biol. 2018, vol. 78, no. 2, pp. 389-390*