

# Predicting the effects of the single nucleotide polymorphism A122V on CXC chemokine receptor type 1 of *Bos taurus* (Artiodactyla: Bovidae) cattle by *in silico* analyses

Anete Ferraz Guzzi <sup>1</sup>

Felipe Santos de Luna Oliveira <sup>2</sup>

Márcia Maria de Souza Amaro <sup>2</sup>

Paulo Fernando Tavares Filho <sup>3</sup>

Jane Eyre Gabriel <sup>1,4\*</sup>

<sup>1</sup> Colegiado de Pós-Graduação em Ciências Veterinárias no Semiárido, Campus de Ciências Agrárias  
Universidade Federal do Vale do São Francisco, CEP 56.300-000, Petrolina – PE, Brasil

<sup>2</sup> Colegiado Acadêmico de Ciências Biológicas, Campus de Ciências Agrárias

<sup>3</sup> Colegiado Acadêmico de Engenharia da Computação, Campus de Juazeiro  
Universidade Federal do Vale do São Francisco, CEP 48902-300, Juazeiro – BA, Brasil

<sup>4</sup> Colegiado Acadêmico de Medicina, Campus Sede  
Universidade Federal do Vale do São Francisco, CEP 56304-917, Petrolina – PE, Brasil

\* Autor para correspondência  
jane.gabriel@univasf.edu.br

Submetido em 30/06/2017

Aceito para publicação em 11/10/2017

## Resumo

**Predizendo os efeitos do polimorfismo de nucleotídeo único A122V sobre o receptor de quimiocina CXC do tipo 1 de bovino *Bos taurus* (Artiodactyla: Bovidae) por análises *in silico*.** Este estudo objetivou prever bioquimicamente as estruturas primárias e secundárias da proteína receptora de quimiocina CXCR1 bovina não polimórfica e polimórfica carregando o polimorfismo A122V por análises *in silico*. Duas sequências da proteína CXCR1 de bovino *Bos taurus* foram selecionadas a partir da base de dados de sequências de proteínas UniProtKB/Swiss-Prot: a) uma sequência não polimórfica (A7KWG0), contendo o aminoácido alanina (A) na posição 122, e b) uma sequência polimórfica, apresentando o aminoácido valina (V) na mesma posição. As estruturas primárias e secundárias da proteína foram preditas empregando os programas ProtParam e Chou & Fasman Protein Secondary Structure Prediction CFSSP. Diferenças nos parâmetros físicos e químicos não foram previstas a partir da estrutura primária de ambas as proteínas analisadas. A presença de um domínio helicoidal, situado nas posições 100 e 150, foi exclusivamente encontrada na proteína CXCR1 não polimórfica. Resíduos de aminoácidos de propriedades bioquímicas variáveis foram detectados na posição 122 na proteína CXCR1 dos ruminantes e humana, sugerindo que esse peptídeo é altamente polimórfico em vertebrados. Os resultados aqui descritos predizem diferenças no padrão da estrutura secundária da proteína CXCR1 bovina não polimórfica e polimórfica A122V empregando ferramentas de Bioinformática.

**Palavras-chave:** Bioinformática; Conservação proteica; Estrutura primária e secundária; Proteína CXCR1

## Abstract

This study aimed to perform *in silico* analyses of the primary and secondary structures of the bovine chemokine receptor CXCR1, the non-polymorphic and A122V-harboring polymorphic proteins to predict differences. Two sequences of the CXCR1 protein of *Bos taurus* cattle were selected from the non-redundant protein sequence database UniProtKB/Swiss-Prot: a) a non-polymorphic sequence (A7KWG0), with alanine (A) at position 122, and b) a sequence harboring the causal polymorphism A122V with substitution by valine (V) at the same position. Protein primary and secondary structures were analyzed using the ProtParam program and Chou & Fasman Protein Secondary Structure Prediction CFSSP algorithm. No differences in physical or chemical parameters were predicted from the primary structure of the two bovine protein sequences. The presence of a helix domain situated between positions 100 and 150 was only found in the non-polymorphic CXCR1 protein. Amino acid residues with different biochemical features were detected at position 122 in the ruminant and human CXCR1 protein sequences, suggesting that this peptide seems to be highly polymorphic in vertebrates. Findings described herein predict differences in the secondary structure pattern of non-polymorphic and polymorphic A122V-harboring CXCR1 proteins using bioinformatics tools.

**Key words:** Bioinformatics; CXCR1 protein; Primary and secondary structure; Protein conservation

## Introduction

The CXC chemokine receptors are integral membrane proteins that specifically bind and respond to cytokines of the CXC chemokine family. There are currently seven known CXC chemokine receptors in mammals, named CXCR1 through CXCR7. The CXC chemokine receptor type 1 (CXCR1) belongs to a class A, rhodopsin-like G-protein-coupled receptor (GPCR), the largest class of integral membrane proteins responsible for cellular signal transduction (TENSEN et al., 1999; BOOTH et al., 2002). This transmembrane protein is one of two high-affinity receptors for the CXC chemokine interleukin-8 (IL-8), a major mediator of immune and inflammatory responses implicated in several disorders. Since IL-8 released in response to inflammatory stimuli binds to the extracellular side of CXCR1, such ligand-activated intercellular signaling pathways result in the neutrophil migration to the site of inflammation (FUTOSI et al., 2013).

Despite its importance, the molecular mechanism of CXCR1 protein is poorly understood due to the limited structural information available. Over the past years, the three-dimensional structure of human CXCR1 was characterized in liquid crystalline phospholipid bilayers by a consensus fold of a G-protein-coupled receptor (GPCR), with seven transmembrane helices (TM1-TM7) connected by three extracellular loops (ECL1-ECL3) and three intracellular loops (ICL1-

ICL3) (PARK et al., 2012). The structure of CXCR1 shares significant similarities with that of CXCR4, the only other chemokine receptor whose structure has been determined. According to Park et al. (2012), in both CXCR1 and CXCR4, charged residues are mainly located near the membrane-water interface, with negative charges clustered in the extracellular loops where they can play a role in ligand binding and receptor activation. Moreover, four charged residues, contributed by helices TM2 (Asp85), TM3 (Lys117) and TM7 (Asp288, Glu291), form a polar cluster in the core of the helical bundle of CXCR1 that may have important consequences for ligand binding and receptor signal transduction (PARK et al., 2012).

Over the past decade, some reports in the literature have described a strong association between polymorphisms in the bovine CXCR1 gene and pathogen-specific incidence rate of clinical mastitis (VERBEKE et al., 2014; PAWLIK et al., 2015; BAGHERI et al., 2016). Recently, Pokorska et al. (2016) identified single nucleotide polymorphisms located within the coding region and the 3'UTR site of the bovine CXCR1 gene that were significantly associated with the susceptibility of cows to clinical mastitis. Interestingly, of the identified mutations, only the SNP at position 365 relative to the start codon (c.+365T>C), characterized as a nonsynonymous mutation, induced a change in the coded protein [GCC (Ala) to GTC (Val)

at the 122nd amino acid]. Thus, to determine if the replacement of this amino acid can result in structural changes in this transmembrane receptor, the current study aimed to characterize biochemically non-polymorphic and polymorphic A122V-harboring CXCR1 protein sequences of *B. taurus* by predicting differences due to the primary and secondary structures of these peptides using computational biology tools.

## Material and Methods

### Determination of the primary and secondary structures of non-polymorphic and polymorphic A122V-harboring CXCR1 protein sequences of cattle

Two sequences of CXCR1 protein of *B. taurus* were selected from the non-redundant protein sequence database UniProtKB/Swiss-Prot (BOUTET et al., 2007): a) a non-polymorphic sequence (A7KWG0), containing the alanine (A) at position 122, and b) a polymorphic sequence harboring the causal polymorphism A122V with the substitution by valine (V) at the same position. Analyses of the primary and secondary structures of the amino acid residues of both protein sequences were carried out using the ProtParam program and Chou & Fasman Protein Secondary Structure Prediction CFSSP algorithm, as described in detail by Costa et al. (2013) and Costa et al. (2018). Theoretical isoelectric point (pI), estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) were some of the biochemical parameters determined under these experimental conditions. Moreover, comparative data of structural components (sheet, helix, turn and coil) typical of protein secondary structure were also analyzed for the selected non-polymorphic and polymorphic protein sequences. Alternatively, amino acid residues at position 122 of the CXCR1 protein sequence of humans (*Homo sapiens*, P25024, 350 residues), *B.*

*taurus* cattle (A7KWG0, 360 residues), *Bos indicus* cattle (A0A023JCG5, 360 residues) and sheep (*Ovis aires*, W5QDY9, 368 residues) were directly compared and identified to assess their conservation in these animals.

## Results

No differences in the physical and chemical parameters were predicted from primary structure analyses between non-polymorphic and polymorphic A122V-harboring protein sequences of bovine CXCR1. Data of the primary structure of the non-polymorphic and polymorphic sequences demonstrated a theoretical pI of 8.94 and molecular weight of approximately 41 kDa, with a total number of negatively (asparagine+glutamic acid) and positively (arginine+lysine) charged residues of 26 and 33, respectively. Similar values for instability index (32.49) and GRAVY (0.508) were found for the two proteins analyzed. Interestingly, half-life estimates inferred from *in vitro* mammalian reticulocytes demonstrated identical values of 30 hours for non-polymorphic and polymorphic CXCR1.

Secondary structure of non-polymorphic and polymorphic A122V-harboring CXCR1 demonstrated a marked difference, with the presence of a helix domain between residues 100 and 150 only in the non-polymorphic protein (Figure 1). This domain was absent at that position in the polymorphic CXCR1 protein (Figure 1).

Direct comparison of the amino acid residues identified at position 122 in the CXCR1 protein sequence of humans, *B. taurus* (non-polymorphic), *B. indicus* and sheep demonstrated the presence of hydrophobic and nonpolar amino acid residues with variable biochemical and physical proprieties, such as tyrosine (Y) and phenylalanine (F) with an aromatic ring and alanine (A) and valine (V) with an aliphatic R-group (arrow in Figure 2).

FIGURE 1: Prediction of the secondary structure of the non-polymorphic (A) and polymorphic A122V-harboring (B) CXCR1 proteins of *Bos taurus* cattle. The arrow indicates the presence of a helix domain in only the non-polymorphic CXCR1 protein sequence between positions 100 and 150. The possible structural features are presented as colored vertical bars: helix in red, sheet in green, turn in blue and coil in yellow.

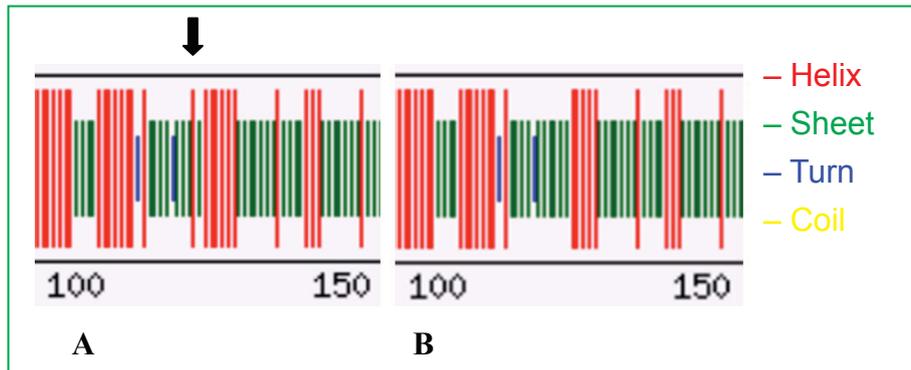


FIGURE 2: Direct comparison of the amino acid residues identified at position 122 in the CXCR1 protein sequences of human (P25024), bovine (*B. taurus*, A7KMG0), bovine (*B. indicus*, A0A023JCG5) and sheep (*Ovis aires*, W5QDY9). Arrow indicates all residues at position 122 in the sequences of interest. Amino acids Y: tyrosine, A: alanine, V: valine and F: phenylalanine.

P25024 CXCR1_HUMAN	100	V	N	G	W	I	F	G	T	F	L	C	K	V	V	S	L	L	K	E	V	N	F	<b>Y</b>	S	G	I	L	L
A7KMG0_BOVIN	109	-	-	-	-	-	-	-	-	-	A	K	G	W	I	F	G	T	P	L	C	K	V	<b>A</b>	S	L	L	K	E
A0A023JCG5_BOSIN	109	-	-	-	-	-	-	-	-	-	A	K	G	W	V	F	G	T	P	L	C	K	V	<b>V</b>	S	L	L	K	E
W5QDY9_SHEEP	117	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	<b>F</b>	G	T	P	L	C

## Discussion

Comparative biochemical parameters of the primary structure of the non-polymorphic and polymorphic CXCR1 proteins predicted a long half-life of 30 hours for both proteins. Biological half-time is understood to be the time for a material or substance to lose one-half of its quantity in the environment or that of its concentration in a particular body organ (NELSON; COX, 2005). A long half-life of 19 hours for rat and human CXCR1 protein was also reported by Garau et al. (2006) by *in vivo* analyses. Thus, the high half-life value for non-polymorphic and polymorphic A122V-harboring CXCR1 proteins demonstrated under these experimental conditions seems to indicate the strong ability of this peptide to exert its biological effects more extensively in the recruitment of neutrophils during inflammation processes. Accordingly, the instability index and GRAVY data for the non-polymorphic and polymorphic CXCR1 proteins support evidence of the great structural stability of the proteins tested.

The secondary structure pattern of the two bovine CXCR1 proteins revealed the presence of a helix domain situated between positions 100 and 150 only in the non-polymorphic protein (Figure 1). The alpha helix is a common motif in the secondary structure of proteins and is a right hand-coiled or spiral conformation in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid located three or four residues earlier along the protein sequence (ROST, 2001). Among different motifs associated with mature spatial structure in the proteins, the alpha helix is the most regular and the most predictable from the sequence, as well as the most prevalent. Different amino-acid sequences have different propensities for forming the alpha-helical structure, since methionine, alanine, leucine, and uncharged glutamate and lysine all have especially high helix-forming propensities, whereas proline and glycine have poor helix-forming propensities (PADMANABHAN; BALDWIN, 1991; PACE; SCHOLTZ, 1998). Helices are also the most

common element in the protein structure that crosses biological membranes in the transmembrane proteins (KUROCHKINA, 2010). Thus, the helical structure can satisfy all backbone hydrogen-bonds internally, leaving no polar groups exposed to the membrane if the side chains are hydrophobic. Within this context, the replacement of alanine with high helix-forming propensity by valine with limited helix-forming propensity in the CXCR1 protein sequence harboring the A122V polymorphism could explain the absence of a helix domain in the secondary structure of the polymorphic protein (Figure 1).

Some reports in the literature have described that single nucleotide polymorphisms result in impaired membrane localization and reduced ATPase activity in the multidrug ATP-binding cassette membrane transporter ABCG2, which consists of six transmembrane regions and one ATP-binding cassette (MIZUARAI et al., 2004). Based on the three-dimensional structure of the human CXCR1 protein reported in detail by Park et al. (2012), the presence of the alanine and valine residues at position 122 in the non-polymorphic and polymorphic CXCR1 proteins, respectively, seems to be very far away from regions of the protein involved in its function (binding sites, active sites). Although substituting one amino acid for another with very dissimilar properties (nonpolar, hydrophobic amino acids) is less likely to cause functional protein alterations, the replacement of the small amino acid alanine by the large amino acid valine could sterically hinder packing during protein folding. In the meantime, theoretical studies have shown that a silent mutation in a complex, mammalian membrane transport protein alters substrate specificity, affecting the timing of cotranslational folding and resulting in altered function of consequence to clinically important aspects, as postulated by Kimchi-Sarfaty et al. (2007). Accordingly, we can infer that the substitution of alanine by valine in the polymorphic CXCR1 protein could directly affect post-translational folding mechanisms of this transmembrane receptor, consequently leading to higher susceptibility of bovine animals to clinical mastitis.

Interestingly, the direct comparison of the amino acid residues identified at position 122 in the CXCR1

protein sequences between the different animals selected demonstrated the presence of residues with variable biochemical and physical properties (Figure 2). Such evidence may suggest that this transmembrane protein seems to be highly polymorphic in the vertebrates analyzed. In fact, Pighetti et al. (2012), investigating haplotypes or the combination of single nucleotide polymorphisms on a single bovine allele, found that such modifications were capable of causing amino acid changes and increasing the specificity of an animal's genetic background. These authors concluded that the bovine CXCR1 gene is highly polymorphic with potential implications towards genetic selection and understanding host factors that increase the risk of infection (PIGHETTI et al., 2012).

In conclusion, the findings described herein demonstrated the prediction of differences in the structure-folding pattern of the non-polymorphic and polymorphic CXCR1 proteins of cattle using bioinformatics tools. Accordingly, the presence of a different motif in the secondary structure of the CXCR1 protein harboring the A122V polymorphism may exert a negative effect on its biological action during inflammatory processes, supporting practical evidence that this polymorphism is associated with the susceptibility of cows to clinical mastitis.

## Acknowledgments

This study was supported by the Bioinformatics and Computational Biology Group "BIO in BYTES." Dr. A. Leyva (USA) provided English editing of the manuscript.

## References

- BAGHERI, M.; MORADI-SHARHRBABA, M.; MIRAIE-ASHTIANI, R.; SAFDARI-SHAHROUDI, M.; ABDOLLAHI-ARPAHAHI, R. Case-control approach application for finding a relationship between candidate genes and clinical mastitis in Holstein dairy cattle. *Journal of Applied Genetics*, Poznan, v. 57, p. 107-112, 2016.
- BOOTH, V.; KEIZER, D. W.; KAMPHUIS, M. B.; CLARK-LEWIS, I.; SYKES, B. D. The CXCR3 binding chemokine IP-10/CXCL10: structure and receptor interactions. *Biochemistry*, Washington, v. 41, p. 10418-10425, 2002.

- BOUTET, E.; LIEBERHERR, D.; TOGNOLLI, M.; SCHNEIDER, M.; BAIROCH, A. UniProtKB/Swiss-Prot. **Methods in Molecular Biology**, Berlin, v. 406, p. 89-112, 2007.
- COSTA, F. F. P.; LIDANI, K. C. F.; NUNES, S. L.; GOUVEIA, J. J. S.; GABRIEL, J. E. Phylogenetical clustering among domestic ruminants from diacylglycerol-acyltransferase-1 lipogenic enzyme and biochemical characterization of its single nucleotide polymorphism K232A in bovine *Bos taurus*. **Animal Biology Journal**, New York, v. 4, p. 225-234, 2013.
- COSTA, F. F. P.; NOGUEIRA, J. F.; GOUVEIA, J. J. S.; GABRIEL, J. E. Characterizing the polymorphism K232A of the diacylglycerol-acyltransferase-1 lipogenic enzyme of bovine *Bos taurus* using *in silico* comparative protein prediction analyses. **Brazilian Journal of Biology**, São Carlos, v. 78, n. 2, 2018, in press.
- FUTOSI, K.; FODOR, S.; MÓCSAI, A. Neutrophil cell surface receptors and their intracellular signal transduction pathways. **International Immunopharmacology**, New York, v. 17, p. 638-650, 2013.
- GARAU, A.; BERTINI, R.; MOSCA, M.; BIZZARRI, C.; ANACARDIO, R.; TRIULZI, S.; ALLEGRETTI, M.; GHEZZI, P.; VILLA, P. Development of a systemically-active dual CXCR1/CXCR2 allosteric inhibitor and its efficacy in a model of transient cerebral ischemia in the rat. **European Cytokine Network**, Montrouge, v. 17, p. 35-41, 2006.
- KIMCHI-SARFATY, C.; OH, J. M.; KIM, I. W.; SAUNA, Z. E.; CALCAGNO, A. M.; AMBUDKAR, S. V.; GOTTESMAN, M. M. A "silent" polymorphism in the MDR1 gene changes substrate specificity. **Science**, Washington, v. 15, p. 525-528, 2007.
- KUROCHKINA, N. Helix-helix interactions and their impact on protein motifs and assemblies. **Journal of Theoretical Biology**, Amsterdam, v. 264, p. 585-592, 2010.
- MIZUARAI, S.; AOZASA, N.; KOTANI, H. Single nucleotide polymorphisms result in impaired membrane localization and reduced ATPase activity in multidrug transporter ABCG2. **International Journal of Cancer**, New York, v. 109, p. 238-246, 2004.
- NELSON, L. D.; COX, M. **Lehninger principles of biochemistry**. 4 ed. New York: W. H. Freeman and Company, 2005. 1119 p.
- PACE, C. N.; SCHOLTZ, J. M. A helix propensity scale based on experimental studies of peptides and proteins. **Biophysical Journal**, Cambridge, v. 75, p. 422-427, 1998.
- PADMANABHAN, S.; BALDWIN, R. L. Straight-chain non-polar amino acids are good helix-formers in water. **Journal of Molecular Biology**, Amsterdam, v. 219, p. 135-137, 1991.
- PAWLIK, A.; SENDER, G.; KAPERA, M.; KORWIN-KOSSAKOWSKA, A. Association between interleukin 8 receptor  $\alpha$  gene (CXCR1) and mastitis in dairy cattle. **Central European Journal of Immunology**, Poznan, v. 40, p. 153-158, 2015.
- PARK, S. H.; DAS, B. B.; CASAGRANDE, F.; TIAN, Y.; NOTHNAGEL, H. J.; CHU, M.; KIEFER, H.; MAIER, K.; DE ANGELIS, A. A.; MARASSI, F. M.; OPELLA, S. J. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. **Nature**, London, v. 491, p. 779-783, 2012.
- PIGHETTI, G. M.; KOJIMA, C. J.; WOJAKIEWICZ, L.; RAMBEAUD, M. The bovine CXCR1 gene is highly polymorphic. **Veterinary Immunology and Immunopathology**, New York, v. 145, p. 464-470, 2012.
- POKORSKA, J.; DUSZA, M.; KUŁAJ, D.; ŻUKOWSKI, K.; MAKULSKA, J. Single nucleotide polymorphisms in the CXCR1 gene and its association with clinical mastitis incidence in Polish Holstein-Friesian cows. **Genetics and Molecular Research**, Ribeirão Preto, v. 15, p. 1-8, 2016.
- ROST, B. Review: protein secondary structure prediction continues to rise. **Journal of Structural Biology**, San Diego, v. 134, p. 204-218, 2001.
- TENSEN, C. P.; FLIER, J.; VAN DER RAAIJ-HELMER, E. M.; SAMPAT-SARDJOEPERSAD, S.; VAN DER SCHORS, R. C.; LEURS, R.; SCHEPER, R. J.; BOORSMA, D. M.; WILLEMZE, R. Human IP-9: a keratinocyte derived high affinity CXCR3-chemokine ligand for the IP-10/Mig receptor (CXCR3). **Journal of Investigative Dermatology**, New York, v. 112, p. 716-722, 1999.
- VERBEKE, J.; VAN POUCKE, M.; PEELMAN, L.; PIEPERS, S.; DE VLIEGHER, S. Associations between CXCR1 polymorphisms and pathogen-specific incidence rate of clinical mastitis, test-day somatic cell count, and test-day milk yield. **Journal of Dairy Science**, Champaign, v. 97, n. 12, p. 7927-7939, 2014.