Article Title: Characterization and sequence analysis of cysteine and glycine-rich protein 3 in Egyptian native cattle and river native buffalo cDNA sequences

Author(s): Ahlam A. Abou Mossallam, Nevien M. Sabry, Eman R. Mahfouz*, Mona A. Bibars and Soheir M. El Nahas

Affiliation: Department of Cell Biology, Genetic Engineering Division, National Research Center, Dokki, Giza, Egypt.

Abstract: Cysteine and glycine rich protein, CSRP3 also referred to as the muscle LIM protein (MLP), has been investigated in native Egyptian cattle and buffalo (river buffalo). RNA extraction and cDNA synthesis were conducted from different tissue samples. Primers specific for CSRP3 were designed using known cDNA sequences of Bos taurus published in database with different accession numbers. Polymerase chain reaction (PCR) was performed and products were purified and sequenced. Sequence analysis and alignment were carried out using CLUSTAL W (1.83). Multiple nucleotide sequence alignment between CSRP3 cDNA amplicons of native buffalo and cattle revealed 89% identity. B. taurus CSRP3 mRNA (Cardiac LIM protein) [NM 001024689.2] showed 85 and 87% identity in nucleic acid sequences and 82 and 84% homology in amino acid sequences with native cattle and buffalo, respectively. A 90% homology was detected between the amino acid sequences of river buffalo and native cattle. Forty nine translated amino acids out of 51 in both buffalo and cattle are found to be part of the conserved CSRP3 LIM1 domain which comprises 57 codons. The LIM1 domain in Egyptian buffalo and cattle CSRP3 showed only 87 and 85% similarity with B. taurus CSRP3 LIM1 domain, respectively, which are caused mainly by frame shift mutation resulting from a single nucleotide deletion. Sequence nucleotide alignment of both native buffalo and cattle CSRP3 cDNAs sequences and B. taurus whole genome showed high percent identity (94-100%) with B. taurus chromosome 29 |NC-007330.3|. This confirmed the assignment of CSRP3 to cattle chromosome BTA 29 and allowed the indirect assignment of CSRP3 to river buffalo chromosome BBUsp (the homologue of BTA 29) based on the extensive chromosome homology and conservation between cattle and river buffalo.

Key words: CSRP3, cattle, river buffalo.

INTRODUCTION

Cysteine and glycine rich protein family also known as CSRP protein are encoded by the CSRP genes (Weiskirchen et al., 1995) they have regulatory roles in reproduction and development (Arber et al., 1997). Members of the CSRP family are characterized by the presence of two tandemly arranged LIM (Lin-11, Isl-1, Mac-3) domains linked to short glycine-rich regions (Weiskirchen et al., 1995). ‘LIM domains’ are protein structural domains, comprised of two contiguous zinc finger domains, separated by a two-amino acid residue hydrophobic linker (Kadmas and Beckerle, 2004). LIM-domain proteins have been shown to play roles in cytoskeletal organization, organ development and oncogenesis. They are named after their initial discovery in the proteins Lin11, Isl-1 and Mec-3 (Bach, 2000). The vertebrate CSRP family consists of three isoforms CSRP1, CSRP2, and CSRP3 (Sadler et al., 1992; Weiskirchen and Bister, 1993; Arber et al., 1994; Crawford et al., 1994; Weiskirchen et al., 1995); they are...
translated amino acids of the downstream strand of CSRP3 amplicons are presented in Figure 3. CLUSTAL W (1.83) multiple nucleotide sequence alignment between CSRP3 amplicons of *B. taurus* CSRP3 (Cardiac LIM protein) [NM 001024689.2] and both native cattle and river buffalo showed 85 and 87% identity, respectively. Homology between native buffalo and native cattle was 89% (Figure 4). Point mutations were detected in both native cattle and buffalo CSRP3-cDNA. They both have one deletion between nucleotides number 126 and 127. Native cattle showed 9 substitutions in addition to insertion of one cytosine at nucleotide number 137 and thymine and adenine at nucleotides number 140 and 141, respectively. In river native buffalo CSRP3-cDNA, 9 substitutions and insertion of three adenosines (AAA) at nucleotides number 144, 145 and 146 were detected (Figure 4). Native cattle showed 82% homology in CSRP3 amino acid (aa) sequence with *B. taurus* [NP 001019860.1]. However it showed 90% homology with river native buffalo (Figure 5). 84% homology was detected between the aa sequences of river native buffalo CSRP3-cDNA and *B. taurus* CSRP3 mRNA (NP 001019860.1). Forty nine out of 51 translated amino acids in both river native buffalo and native cattle CSRP3 are part of the CSRP3 LIM1–domain protein (57 codons). They were found to be 93% homologous to each other, whereas they showed only 87 and 85% similarity with *B. taurus* CSRP3 LIM1 domain, respectively (Figure 6). Alignment of native cattle and buffalo CSRP3 cDNAs and cattle whole genome revealed that sequences from 1 to 61 bp (exon II) and from 62 to 135 bp (exon III) showed 100 and 94% (in cattle) and 100 and 97% (in buffalo) identity with two segments of *B. taurus* chromosome 29 (NC_007330.3) (from 26936061 to 26936121 bp) and (from 26940354 to 26940428 bp), respectively.

**RESULTS AND DISCUSSION**

*CSRP3* primers amplified a 180-bp segment in muscle cDNA of native cattle (Figure 1) and in muscle and trachea cDNAs of river native buffalo (Figure 2). CSRP3 muscle cDNA amplicons of native cattle and buffalo were sequenced. The resulting nucleotide sequences (154 bp) have been submitted to GenBank and were given accession numbers GU433602 and GQ231523, respectively.

The reverse complement nucleotide sequences and the translated amino acids of the downstream strand of CSRP3 amplicons were presented in Figure 3. CLUSTAL W (1.83) multiple nucleotide sequence alignment between CSRP3 amplicons of *B. taurus* CSRP3 (Cardiac LIM protein) [NM 001024689.2] and both native cattle and river buffalo showed 85 and 87% identity, respectively. Homology between native buffalo and native cattle was 89% (Figure 4). Point mutations were detected in both native cattle and buffalo CSRP3-cDNA. They both have one deletion between nucleotides number 126 and 127. Native cattle showed 9 substitutions in addition to insertion of one cytosine at nucleotide number 137 and thymine and adenine at nucleotides number 140 and 141, respectively. In river native buffalo CSRP3-cDNA, 9 substitutions and insertion of three adenosines (AAA) at nucleotides number 144, 145 and 146 were detected (Figure 4). Native cattle showed 82% homology in CSRP3 amino acid (aa) sequence with *B. taurus* [NP 001019860.1]. However it showed 90% homology with river native buffalo (Figure 5). 84% homology was detected between the aa sequences of river native buffalo CSRP3-cDNA and *B. taurus* CSRP3 mRNA (NP 001019860.1). Forty nine out of 51 translated amino acids in both river native buffalo and native cattle CSRP3 are part of the CSRP3 LIM1–domain protein (57 codons). They were found to be 93% homologous to each other, whereas they showed only 87 and 85% similarity with *B. taurus* CSRP3 LIM1 domain, respectively (Figure 6). Alignment of native cattle and buffalo CSRP3 cDNAs and cattle whole genome revealed that sequences from 1 to 61 bp (exon II) and from 62 to 135 bp (exon III) showed 100 and 94% (in cattle) and 100 and 97% (in buffalo) identity with two segments of *B. taurus* chromosome 29 (NC_007330.3) (from 26936061 to 26936121 bp) and (from 26940354 to 26940428 bp), respectively.

**MATERIALS AND METHODS**

**RNA Isolation and First-strand cDNA synthesis**

RNA extraction was conducted for tissue samples including muscle, lung, trachea, intestine and liver, from both native cattle and buffalo (Grubor et al., 2004). The lack of contaminating genomic DNA was checked out by monitoring negative polymerase chain reaction (PCR) products in the absence of reverse transcriptase. cDNA synthesis was performed using reverse transcriptase (RT)-PCR Ready-to-go kit (Amersham Biosciences).

**Primer design**

Primers specific for the antimicrobial peptide gene *CSRP3* were designed using *Bos taurus* cDNA sequences published in database with different accession numbers. The sequence of the forward and reverse primers was determined using the software Primer 3 (Marone et al., 2001), http://www.genome.wi.mit.edu. PCR primers were selected on the basis that the 5' and 3' ends span exons I and III. The primers were synthesized by Amersham Pharmacia Biotech.

**Polymerase chain reaction (PCR)**

Amplification reactions (100 μl) contained 5 μl of first-strand cDNAs, 0.2 mM dNTPs, 10 mM Tris, 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin (W/V), 1.25 units *Taq* polymerase and 1 μM forward and reverse primers. The reaction mixture was overlaid with sterile mineral oil. PCR was performed using MJ research PTC-100 thermocycler using 1 cycle (3 min.) at 94°C, followed by 30 cycles for 1 min at 94°C, 2 min at 56.5°C, and 2 min at 72°C and finally 1 cycle (10 min.) at 72°C. The reaction products were electro-pheroresed on 1.5% agarose in 1× tris acetate buffer (TAE) containing 0.8 μl of 10 mg/ml ethidium bromide. Primer sequences, annealing temperature, product size and accession number are shown in (Table 1).

**Sequence analysis**

The PCR products were purified and sequenced at the Center of Genetic Engineering, Ain Shams University, Cairo, Egypt. Sequence analysis and nucleotide and amino acid sequence alignments were carried out using CLUSTAL W analysis (Gasteiger et al., 2003) (http://www.ebi.ac.uk/tool/clustalw/index.html). Frame translation of nucleotide sequence was carried out using http://searchlauncher bcm.tmc.edu/seq-util/Options/sixframe.html as shown in Table 1.

**Table 1. DNA sequence of the primers tested.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Accession no.</th>
<th>Ann. Temp.</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cysteine and glycine-rich protein 3 (CSRP3)</td>
<td>TACCAGCGAGAAGATCCAG</td>
<td>NM_001024689</td>
<td>56.5</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>CAGCGCCTTGCCATACC</td>
<td>GI 67010008</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

encoded by an approximately 20-kb genomic sequence with 6 exons (Wang et al., 2006) and contain 192–194 amino acid residues. CSRP1 is expressed in non-striated muscle tissues such as smooth muscle, CSRP2 is expressed in vascular tissue and fibroblasts (Louis et al., 1997), whereas CSRP3 maintains the stability of the contractile apparatus (Arber et al., 1997; Pomies et al., 1997).
Cysteine and glycine-rich protein 3 (CSRP3, cardiac LIM-domain protein) gene encodes a member of the CSRP family of LIM domain proteins, which may be involved in regulatory processes important for development and cellular differentiation. The LIM/double zinc-finger motif in this protein is found in a group of proteins with critical functions in gene regulation, cell growth, and somatic differentiation (Newman et al., 2005).

The 2 LIM domains in CSRP3 of B. taurus and in many other mammalian species lies from codon 10 to 66 and from 120 to 176 (Strausberg et al., 2002). The amplified segments investigated in this study in both native cattle and buffalo comprise 51 codons within LIM1 domain.
Figure 3. The reverse complement nucleotide sequences of the downstream strand of CSRP3 amplified fragment of native cattle (CTN) and river buffalo (BU) trachea cDNAs, and their aa translations.

<table>
<thead>
<tr>
<th>SeqA Name</th>
<th>Len(nt)</th>
<th>SeqB Name</th>
<th>Len(nt)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 BU-CSRP3</td>
<td>154</td>
<td>2 CTN-CSRP3</td>
<td>154</td>
<td>89</td>
</tr>
<tr>
<td>1 BU-CSRP3</td>
<td>154</td>
<td>3 Bos-CSRP3</td>
<td>297</td>
<td>87</td>
</tr>
<tr>
<td>2 CTN-CSRP3</td>
<td>154</td>
<td>3 Bos-CSRP3</td>
<td>297</td>
<td>85</td>
</tr>
</tbody>
</table>

CLUSTAL W (1.83) multiple sequence alignment

Figure 4. CLUSTAL W (1.83) multiple sequence alignment of native cattle (CTN), river buffalo (BU) and Bos taurus (NM_001024689.2) CSRP3. Primer sequences are indicated in underlined bold typeface.
<table>
<thead>
<tr>
<th>SeqA Name</th>
<th>Len(aa)</th>
<th>SeqB Name</th>
<th>Len(aa)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buf-CSRP3</td>
<td>51</td>
<td>Native cattle</td>
<td>51</td>
<td>90</td>
</tr>
<tr>
<td>Buf-CSRP3</td>
<td>51</td>
<td>Bostaurus-CSRP3</td>
<td>70</td>
<td>84</td>
</tr>
<tr>
<td>Native cattle</td>
<td>51</td>
<td>Bostaurus-CSRP3</td>
<td>70</td>
<td>82</td>
</tr>
</tbody>
</table>

Buf-CSRP3  
-----------------  
YHAEEIQCNGRSFHKTCFHCMA CRC RALDSTTVAA HESEIYCKV 43

Native-cattle  
-----------------  
YHAEEIQCNGRSFHKTCFHCMA CRC RALDSTTVAA HESEIYCK V 43

Bos taurus-CSRP3  
MPNWGGGAKCGACEKT VYHAEEIQCNGRSFHKTCFHCMA CRC RALDSTTVAA HESEIYCKV 60

Figure 5. CLUSTAL W (1.83) pairwise comparison of aa sequences of river buffalo (Buf)-CSRP3 cDNA aa, Bos taurus-CSRP3, mRNA[NP_001019860.1 and cattle native (CTN)-CSRP3. Id= Identical ***, Semi-Con= semi-conserved substitutions **, blank indicates no match *"*. LIM-domain protein is in bold typeface.

<table>
<thead>
<tr>
<th>SeqA Name</th>
<th>Len(aa)</th>
<th>SeqB Name</th>
<th>Len(aa)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bos</td>
<td>57</td>
<td>BU</td>
<td>49</td>
<td>87</td>
</tr>
<tr>
<td>Bos</td>
<td>57</td>
<td>CTN</td>
<td>49</td>
<td>85</td>
</tr>
<tr>
<td>BU</td>
<td>49</td>
<td>CTN</td>
<td>49</td>
<td>93</td>
</tr>
</tbody>
</table>

CLUSTAL W (1.83) multiple sequence alignment

Bos  
CGACEKT VYHAEEIQCNGRSFHKTCFHCMA CRC RALDSTTVAA HESEIYCKV CYGRRY 57
BU  
-----------------  
YHAEEIQCNGRSFHKTCFHCMA CRC RALDSTTVAA HESEIYCKV ATS LVRN 49
CTN  
-----------------  
YHAEEIQCNGRSFHKTCFHCMA CRC RALDSTTVAA HESEIYCKV CATS LVR 49

Figure 6. CLUSTAL W (1.83) pairwise comparison of CSRP3-LIM1 domain of Bos taurus [NP_001019860.1], native buffalo (Bu), cattle native (CTN). Id= Identical ***, Semi-Con= semi-conserved substitutions **, blank indicates no match *"*. (from 10-66). LIM1 is a functionally important domain, which is responsible for interaction with α-actinin, an actin-binding protein with multiple roles in different cell types, and with certain muscle-specific transcription factors. In human CSRP3 gene, all mutations predicted an amino acid exchange at highly conserved residues in the functionally important LIM1 domain. Protein-binding studies indicate that mutations in the CSRP3 gene lead to a decreased binding activity of MLP to α-actinin (Geier et al., 2003).

The LIM1 domain of Egyptian buffalo and cattle CSRP3 showed 87 and 85% similarity with B. taurus CSRP3, respectively, which indicates that their CSRP3 amino acid sequences are likely to share functional domains with B. taurus CSRP3 amino acid sequence. However, the percentage similarity reported here between both native cattle and buffalo and B. taurus is rather lower than expected since LIM1 domain is highly conserved. LIM1 domain was found to be 100% conserved between B. taurus (NP_001019860.1), Homo sapiens (NP_003467.1), Mus musculus (NP_038836.1), and Canis lupus familiaris (XP_865543.1). In Sus scrofa gb (ACL82864.1), 98% similarity was reported where amino acid at position 5 changed from E to D. The latter also occurred in addition to a second substitution at position 56 from R to K in Rattus norvegicus gb (EDM07231.1) bringing similarity to 96%. The lower LIM domain similarity between both native cattle and buffalo and B. taurus are
attributed to a nucleotide deletion in both species after nt 126 causing a frame shift mutation. Frame shifts lead to dramatic change of amino acid sequence (Galvani and Slatkin, 2003). Mutations are the source of new variations important for evolution. It creates variations in the gene pool, and the less favorable (or deleterious) mutations are reduced in frequency in the gene pool by natural selection, while more favorable (beneficial or advantageous) mutations tend to accumulate, resulting in evolutionary change (Knight et al., 2006). Beneficial mutations lead to new versions of proteins that help an organism and its future generations better adapt to changes in their environment (Knight et al., 2006). The presence of the frame shift in both native cattle and native buffalo and the higher similarity in LIM1-domain protein occurring between them (93%) may reflect adaptive changes to cope with the harsh environment they are raised in. Further investigations may be needed.

**Figure 7.** BLAST sequence alignment between native cattle (CTN)-CSRP3 and *Bos taurus* chromosome29 (NC_007330.3).

**Figure 8.** BLAST sequence alignment between river buffalo (Bu)-CSRP3 and *Bos taurus* chromosome29 (NC_007330.3).
Chromosomal assignment

Alignment of buffalo and cattle CSRP3 cDNA sequence with *B. taurus* whole genome showed high percent identity. It confirmed the assignment of CSRP3 to cattle chromosome BTA 29 using radiation hybrid (RH) mapping method by Band et al. (2000). Based on the genetic conservation between cattle and river buffalo (El Nahas et al., 2001 and Di Meo et al., 2008), CSRP3 can be assigned to river buffalo chromosome BBU5p, the homologue of BTA 29. CSRP3 gene is genetically conserved since it was mapped to human chromosome 11 (11p15.1) (Fung et al., 1995), and was found to reside on homologous mouse chromosome 7 (Mahy et al., 2002) and rat chromosome 1 (Harhay and Keene, 2003).

ACKNOWLEDGEMENTS

This study was in part supported by the Executive Program on Scientific and Technological Cooperation between the Academy of Scientific Research and Technology/Egypt and the Hungarian Academy of Sciences.

REFERENCES


