IOSR Journal Of Pharmacy
(e)-ISSN: 2250-3013, (p)-ISSN: 2319-4219
Www.Iosrphr.Org Volume 3, Issue 11 (December 2013), Pp 27-30

Functional assignment to *Mus musculus* endostatin mRNA *via* computational analysis

Kheirollah Yari[1], Hassan Rasouli[1, 2*], Mohammad Taher Moradi[1], Maryam sadat Soltani[3]

[1] *Medical Biology Research Center, Kermanshah University of Medical Sciences, P.O. Box 6714869914, Sorkheh lizheh, Kermanshah, Iran.*
[2] *Agronomy and Plant Breeding and Biotechnology Department (APBBD), College of Agriculture and Natural Resources, Razi University, P.O. Box 6715685438 Kermanshah, Iran.*
[3] *Department of Biology,Science and Research Branch, Islamic Azad university,Tehran,Iran*

ABSTRACT: *Endostatin is an endogenous inhibitor of angiogenesis. It was first found secreted in the media of no-metastasizing mouse cells from a hemangioendothelioma cell line in 1997 and was subsequently found in humans. The sequence of endostatin gene (AC: AF257775) has been downloaded from NCBI database and via Conserved Domain Database (CDD), DNA Motif Searching Database (DMSD), DNA binding-site database predicator, the similar domains, motif, DNA binding-site, has been recognized. The result showed in this sequence important motif includes: EGF_1, CTCK_1, and VWFC_1 and 2FE2S_FER_1. Our result showed in this study a total of 10 domain structures which were similar with endostatin protein sequence has been identified.*

KEYWORDS*: endostatin, Mus musculus, computational analysis.*

I. BACKGROUND

Among many anti-antigenic proteins, endostatin was initially isolated from the conditioned media of a Mouse Hemangioendothelioma Endothelial Cells (EOMA Line) by Folkman [Huang *et al.*, 2001; Deininger *et al*., 2003]. Endostatin is a 20-kDa polypeptide that produced through enzymatic cleavage from COOH-terminal of multiplexin a1-collagen type XVIII [Sorensen *et al*., 2002]. Bioinformatics is an interdisciplinary research area, which may be defined as the interface between biological and computational sciences. It greatly helps in management of complex and scattered biological data, sequence analysis, algorithmic designing. However, by using the *in silico* analysis we can analyze the genomic sequences [Pevzner and Shamir, 2011]. Our analysis highlights to the evaluation endostatin gene sequence *via* in silico analysis.

II. METHODOLOGY

Download sequence:

Download the complete sequence for endostatin gene (with accession number: AF257775) in the NCBI database.

Identification of similar domain structure:

To identify similar domain structure, we used the Conserved domain database (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

Identification of the motifs in endostatin gene:

To analyze the nucleotide sequence we searched for the motifs and the motif search software (http://www.genome.jp/tools/motif) was used to identify the motifs in nucleotide sequence.

Identification of DNA binding-site:

To identify DNA binding-site, we used the BindN database (http://bioinfo.ggc.org/cgi-bin/bindn).

III. RESULTS AND DISCUSSION

Motif structure:

In silico analysis showed the motif structure for this gene includes (Table 1 and Fig. 1):

a) 2FE-2S_FER_1 Motif, or ferredoxins, iron-sulfur binding region signature. Motif 2Fe-2S is a structural motif, from the comparison of the coding proteins between rice and spinach chlorine monooxygenase (CMOs), rice CMO potentially shares two conservative motifs including a Rieske-type [2Fe-2S] (Marwa *et al*., 2011) cluster and a mononuclear non-heme Fe binding sequence. These motifs are considered to be essential for the function of CMO (Gray *et al*., 1997).

b) VWFC_1 Motif. VWFC motif has conserved cysteine which was found in many cereal crops like rice (Fan *et al*., 2006).
c) EGF_1 Motif. The EGF_1 motif or EGF-like domain signature 1 is an evolutionary conserved protein domain, which derives its name from the epidermal growth factor where it was first described. It comprises about 30 to 40 amino-acid residues and has been found in a large number of mostly animal proteins (Downing, 1996).
d) CTCK_1 Motif or C-terminal cystine knot signature. The cystine-knot motif, made up of three intertwined disulfide bridges, is a unique feature of several toxins, cyclotides and growth factors, and occurs in a variety of species, including fungi, insects, molluscs and mammals (Shalini, 2011)

Table 1. Type of identified motif structure in the endostatin gene

Found Motif	Position	PROSITE	Description	symbol
EGF_1	132..143 241..252 333..344 361..372	PS00022	EGF-like domain signature 1.	
CTCK_1	255..293	PS01185	C-terminal cystine knot signature.	
VWFC_1	1..55 446..501 447..501	PS01208	VWFC domain signature.	
2FE2S_FER_1	155..163	PS00197	2Fe-2S ferredoxin-type iron-sulfur binding region signature.	

Fig. 1: Map of the motif structure sequence in the endostatin gene in *Mus musculus*.

IV. SIMILAR DOMAIN STRUCTURE:

Our result showed in this study a total of 10 domain structures which were similar with endostatin protein sequence has been identified (Fig 2). This domain includes: collagen alpha 1(XVIII) chain – human, type XVIII collagen (*Bilateria*, *Taxonomy ID:* 33213), type XVIII collagen long variant (*Euteleostomi*, *Taxonomy ID:* 117571), type XV collagen (*Eutheria*, *Taxonomy ID:* 9347), type XVIII collagen (*Eutheria*, *Taxonomy ID:* 9347), collagen type XVIII, alpha 1 (*Euteleostomi*, *Taxonomy ID:* 117571), alpha-1(XVIII) collagen (*Eutheria*, *Taxonomy ID:* 9347), CLE-1A protein (*Caenorhabditis*, *Taxonomy ID:* 6237), alpha-1(XVIII) collagen (*Murinae*, *Taxonomy ID:* 39107) and collagen alpha-1(XVIII) (*Euteleostomi*, *Taxonomy ID:* 117571).

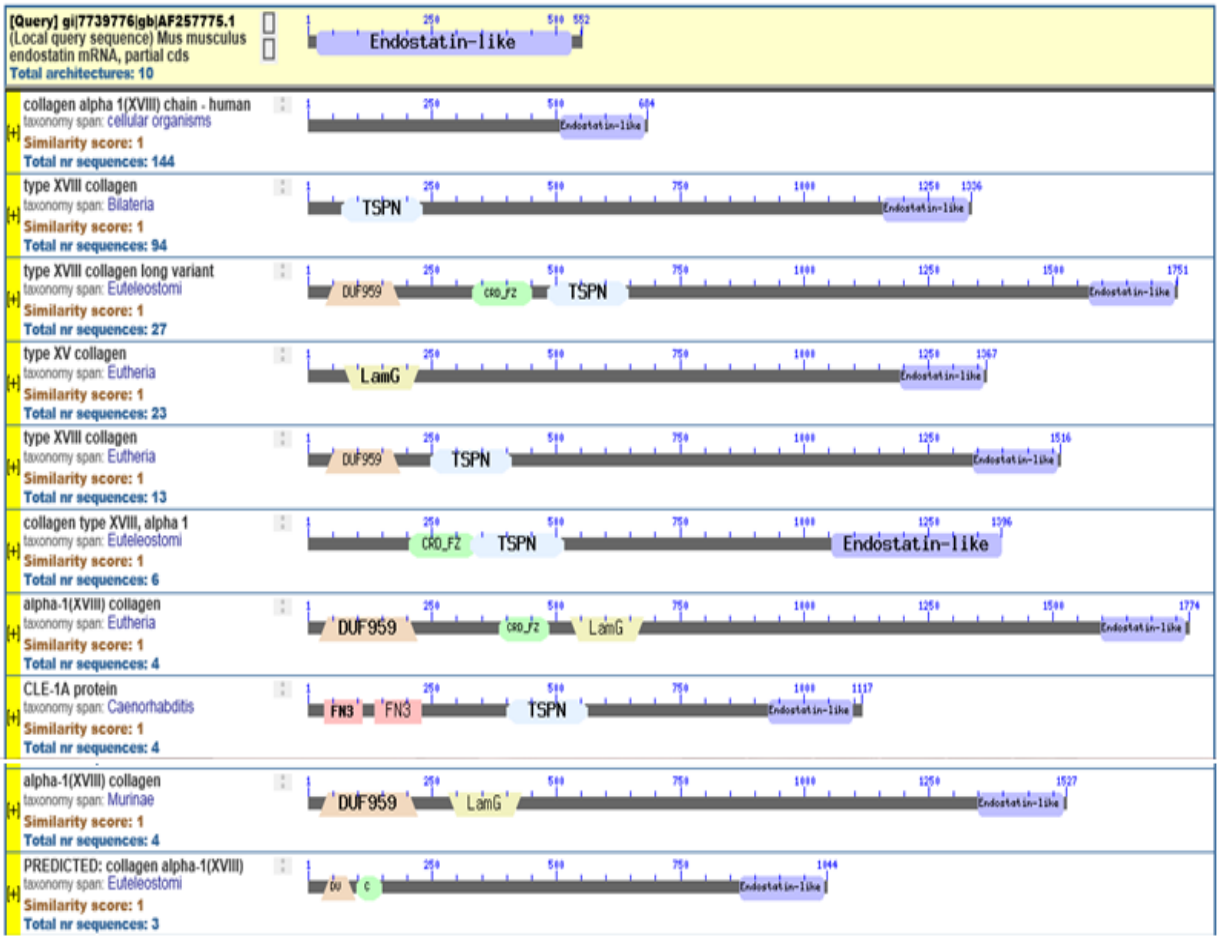

Fig. 2: similar domain structure

Identification of DNA binding-site:

To identify DNA binding-site, we used the BindN database. Our result showed in the endostatin protein sequence have been a total of 43 DNA binding-site (Fig.3)

```
Sequence:    HTHQDFQPVLHLVALNTPLSGGMRGIRGADFQCFQQARAVGLSGTFRAFLSSRLQDLYSI
Prediction:  --------------------+---+--+----------+----+-+-+---+-+-------
Confidence:  32337878997988742245435646767686654397743553859225528723725 6

Sequence:    VRRADRGSVPIVNLKDEVLSPSWDSLFSGSQGQLQPGARIFSFDGRDVLRHPAWPQKSVW
Prediction:  -++--+-+-------------+--+----+--------+--+-------+------++--
Confidence:  94665875758969284782244465523525253348684488926875538322445 2

Sequence:    HGSDPSGRRLMESYCETWRTETTGATGQASSLLSGRLLEQKAASCHNSYIVLCIENSFMT
Prediction:  --+--+-++----+--+-+++++--+-+-++--+-+-------+---------------+
Confidence:  26833849944326735285466247455683557788432677332239998974245 8

Sequence:    SFSK
Prediction:  +-++
Confidence:  7268
```

Fig. 3: Binding residue in endostatin protein sequence. Binding residues are labeled with"+" and in red. Non-binding residues are labeled with "-"and in green.

V. DISCUSSION

This computational analysis focuses on identified a number of motifs structure and similar domains and DNA binding site in endostatin gene. *In silico* analysis provides an efficient way to indicating motifs sequence and similar domains and DNA binding site in genome (Zhu, 1996; Lloyd *et al*., 2001). However, by using the *in silico* analysis we can predicating and identified the motifs structure and similar domains and DNA binding site in gene sequence.

REFERENCES

[1] Huang XJ, Wong MKK, Zhao Q, Zhu ZY, Wang KZQ, Huang N, Ye CS, Gorelik E (2001) Li MF (2001) Soluble recombinant endostatin purified from Escherichia coli: antiangiogenic activity and antitumor effect (vol 61, pg 478). Cancer Res 61(15):5956.

[2] Deininger MH, Wybranietz WA, Graepler FTC, Lauer UM, Meyermann R, Schluesener HJ (2003) Endothelial endostatin release is induced by general cell stress and modulated by the itric oxide/cGMP pathway. Faseb J 17(10):1267–1276.

[3] Sorensen DR, Read TA, Porwol T, Olsen BR, Timpl R, Sasaki T, Iversen PO, Benestad HB, Sim BKL, Bjerkvig R (2002) Endostatin reduces vascularization, blood flow, and growth in a rat gliosarcoma. Neuro-Oncology 4(1):1–8

[4] Zhu, J.K. (2002). Salt and drought stress signal transduction in plants. Annu. Rev. Plant. Biol. 53, 247–273.

[5] Lloyd, G., Landini, P., Bushby, S. (2001). Activation and repression of transcription initiation in bacteria. Essays Biochem. 37, 17–31

[6] Pevzner, P., Shamir, R. 2011. Bioinformatics for biologists. Cambridge University Press. 394. p

[7] Shalini, L., Acharya, K. R. (2011). Tying the knot the cystine signature and molecular-recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. FEBS Journal 278, 4304–4322.

[8] Marwa, N. M. E., E-seuodi, A. A., Abd-Elhady, A. A., Edris, S. E. A., Abdelhamid, N. R., Hassan, A. H. M. (2011). Utilization of Some EST Derived SSR Markers to Detect Glutenin Gene Loci in Wheat. Journal of Applied Sciences Research, 7(12): 2147-2155.

[9] Gray, J., Close, P.S., Briggs, S.P., Johal, G.S. (1997). A novel suppressor of cell death in plants encoded by the Lls1 gene of maize. Cell, 89:25-31.

[10] Fan.C.H., Xing, Y.Z., Mao, H.L., Lu, T.T., and Han, B. (2006). GS3, a major *QTL* for grain length and weight and minor *QTL* for grain width and thickness in rice, encodes a putative transmembrane protein. Theor. Appl. Genet., 112: 1164-1171.

[11] Downing, A. K., Knott, V., Werner, J.M, Cardy, C. M., Campbell, I. D., Handford, P. A. (1996). Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders. Cell 85 (4): 597–605.