INNSPUB

International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 4, No. 11, p. 249-256, 2014

RESEARCH PAPER OPEN ACCESS

Sequence and structural analysis of β- actin protein of fishes, using bioinformatics tools and techniques

U.R. Acharya*, S.K. Behera, S.N. Swain, M.K. Panda, A.R. Mistri, B.B. Sahoo

B.I.F.Centre, P.G. Department of Zoology, Berhampur University, India

Key words: β- actin, Homologs, Secondary structure, NCBI, fishes, *Bos Taurus*.

http://dx.doi.org/10.12692/ijb/4.11.249-256 Article published on June 11, 2014

Abstract

Actins are highly conserved proteins that are involved in various types of cell motility and are ubiquitously expressed in all eukaryotic cells. In vertebrates, three main groups of actin isoforms namely alpha, beta, and gamma have been identified. The alpha actins are found in muscles with contractile functions whereas beta and gamma actins co-exist in the cytoskeleton. For establishing evolutionary inter-relationship, sequence analysis and alignment of protein sequences, seems to be crucial. In the present study, a bioinformatics approach has been adopted to explore the sequence, structure and phylogenetic analysis of β-actin protein from eight commercially important fishes and compared with the standard sequences like 1HLU *Bos taurus*, 2OAN *Bos taurus* and 2BTF *Bos taurus*, which were predicted to be the closest homologs for the 11sequences of the β actin protein of 8 fishes. From the comparative analysis it was concluded that 1HLU *Bos taurus* affords highest sequence coverage with low E-value with the 11 sequences of eight fishes under study. Presence of conserved residues in the sequences and the phylogenetic relationship among fishes with respect to the β-actin protein, have been discussed.

* **Corresponding Author:** U.R.Acharya ✉ ura_zl@rediffmail.com

Introduction

Actin was first identified as a part of the protein complex acto-myosin responsible for producing the contractile force in skeletal muscle (Straub, 1942). It is one of the most abundant proteins in human tissues and is widely expressed in nearly all types of eukaryotic cells. In non-muscle cells, especially those that undergo rapid movements or that resist mechanical stresses, actin protein seems to be highly active near the cell surface (Condeelis, 1993, Stossel, 1994). Structurally, actin is a globular protein with a molecular weight 42-kDa (sheterline and Sparrow,1994), densely found in all eukaryotic cells with exception in nematode sperm where it may be present at very low concentrations. It is also one of the highly-conserved proteins, sharing approximately 20% diversity from algae to humans and, hence, participates in many important cellular processes including muscle contraction, cell motility, cell division and cytokinesis, vesicle and organelle movement, cell signaling, and the establishment and maintenance of cell junctions and cell shape (Pollard and Borisy, 2003,Bernstein and Bamburg, 2010,Doi *et.al.*,2010, Makioka *et al.*, 2011).

The corresponding β-actin gene (*actb*) belongs to the actin multigene family (Kislauskis *et al.*, 1997) and takes part in the initiation of transcription (Hofmann *et al.*, 2004). The gene for β-actin has been cloned from both vertebrates and invertebrates and studies indicate that β-actin is constitutively expressed in most eukaryotic non- muscle cells (Quitscke *et al.*,1989) and it has been widely used as an internal control gene in the quantification of mRNA levels (Lee, 2000).

Studies indicate that there are six genes encoding six vertebrate actins that are classified according to where they are predominately expressed. α-skeletal-Actin, α-smooth-actin, α- cardiac-actin and γ-smooth actin are primarily found in muscle cells, whereas cytoplasmic β- cyto-actin and γ- cyto-actin are ubiquitously and highly expressed in non-muscle cells (Herman,1993).

Functionally, the β- actin gene is highly conserved in the animal kingdom and plays an essential role in maintaining cytoskeletal structure, cellular mobility, cell division and contractile processes (Reece,1992). In addition, *actb* is a housekeeping gene and due to its essential role, *actb* mRNA expression has often been used as a versatile invariant control for gene expression studies (Andreassen *et al.*, 2005, Cao *et al.*, 2007).

The molecular biological study of fish is growing rapidly and β-actin is often used as an internal control gene in the quantification of gene expression (Sohn *et al.*, 2001, Wang and Ge, 2003). However, the information about fish β-actin gene and its protein is still limited, so far its sequence analysis, secondary and tertiary structure elucidation are concerned. In contrast, β-actin is usually used as a loading control, in a number of laboratory-based protocols like PCR and Western blotting analysis. Basing on the importance of β-actin protein, a bioinformatics approach has been preferred to focus on its sequence, structure and phylogenetic analysis. Hence, 11 protein sequences from 8 different commercial fishes namely *Labeo rohita, Labeo calbasu , Chanos Chanos , Catla catla, Heteropneustes fossilis, Clarias batracus, Lates calcarifer , Oreochromis mossambicus* were retrieved from NCBI protein database.

Furthermore, studies' pertain to phylogenetic inter-relationship among fishes taking β-actin protein as a standard seems to be highly fragmentary and inconclusive. Hence, realizing the importance of β-actin protein in fishes, the present study was undertaken to investigate the sequence analysis, structure prediction and phylogenetic relationship among eight fishes utilizing standard bioinformatics tools and techniques.

Materials and methods

Sequence Analysis of β- actin proteins

Entrez Protein sequence retrieval tool developed by National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) was used for retrieval of protein sequences of β- actin protein from

8 commercial fishes under study. In order to cross check the sequence analysis results, UniProt search developed by EBI (European Bioinformatics Institute) was performed. Since, the reports on the 3D (Three Dimensional) structure of the β- actin protein of the fishes in the present study found to be scanty in the structural database like Protein Data Bank (PDB) (www.pdb.org), Basic Local Alignment Search Tool Protein (BLASTP) analysis developed by NCBI was utilized to get the closest homologs for the retrieval of proteins.

Multiple Sequence Alignment of β- actin proteins

ClustalX (1.83) programme was used for multiple sequence alignment in order to predict the conserved amino acid residues in the β-actin protein of fishes. Sequences with higher query coverage and low E-value were considered as the standard sequence for analyzing the conserved positions. The retrieved amino acid sequences of the β- actin protein along with the homolog sequences were loaded in the ClustalX tool for multiple sequence alignment**.** Ghostview tool was also used to visualize the multiple sequence analysis results. Moreover, β-actin protein sequences of the fishes taken for the study, were compared with the standard sequences like 1HLU *Bos taurus*, 2OAN *Bos taurus* and 2BTF *Bos taurus*, retrived from PDB and were predicted to be the closest homologs with the fishes.

Phylogenetic Analysis of of β- actin proteins

In order to search the evolutionary relationships among the protein sequences of fishes, cladograms were constructed with the help of tree viewer tool and Phylip. The cladograms are the branching, treelike diagrams in which the endpoints of the branches represent specific species of organisms.

Structural Analysis of β- actin proteins

Secondary structure prediction tools namely CFSSP, GOR, SOPMA, PSIPred listed in ExPASy (Expert Protein Analysis System) (www.expasy.org/tools/) proteomics server were used to predict the secondary structure of these β- actin protein of these 8 commercial fishes.

Prediction of the 3D (3-Dimensional) structure of the proteins under study, was done by Modeller tool and the structures with least DOPE (Discrete Optimized Protein Energy) score were selected as the model for least energy minimization. Furthermore, the 3D-structures were visualized in RasMol and PyMol viewers which were finally confirmed in Ramchandran Plot.

Results and discussion

Different databases like NCBI, Entrez and UniProt Protein sequence analysis have already reported only 11 amino acid sequences of β-actin proteins of 8 commercial fishes under study. For comparison, β-actin protein from 1HLU *Bos taurus*, 2OAN *Bos taurus* and 2BTF *Bos taurus* were predicted and were found to be the closest homologs for the β- actin proteins of the fishes taken for the study by BLASTP analysis.

Table 1. Shows the no. of 2D structures predicted by different secondary structure prediction tools. (C: Coils, S: Beta pleated Sheets, H: Alpha Helix).

	CFSSP			GOR			SOPMA		
	C	S	H	C	S	H	C	S	H
Labeo rohita 4	0	1	2	5	2	0	6	0	0
Labeo rohita 2	11	8	29	32	9	0	20	10	5
Labeo rohita3	22	5	8	87	38	46	53	29	81
Labeo rohita 1	1	0	6	5	2	0	7	0	0
Chanos Chanos	24	35	21	55	31	41	44	29	46
Catla catla	1	0	6	5	2	0	7	0	0
Labeo calbasu	161	93	156	174	78	103	109	84	139
Clarias batracus	101	67	56	118	34	79	77	34	102
Heteropneustes fossilis	134	31	74	108	29	87	70	48	89
Lates calcarifer	123	60	98	134	69	108	106	69	116
Oreochromis mossambicus	135	76	112	171	90	114	124	87	137

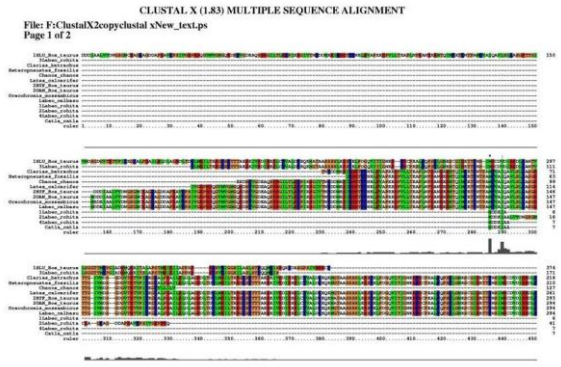

Fig. 1. Clustal X results for 14 β-actin Protein sequences.

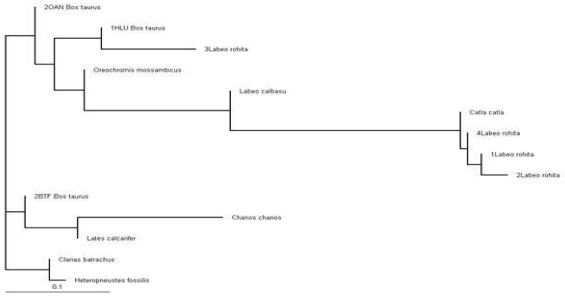

Fig. 2. Phylogenetic tree demonstrating the relatedness of fishes with respect to β-actin protein.

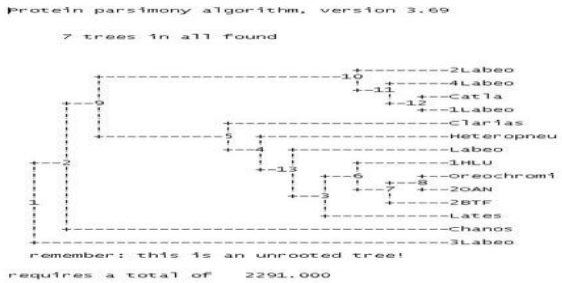

Fig. 3. PHYLIP analysis of β-actin protein from fishes compared with *Bos taurus*.

The amino acid sequence of β-actin protein from *Bos taurus* with a length of 375 amino acid residues was considered as standard sequence for counting the conserved positions of the residues in multiple sequence alignment work, because its 3D structure was reported in PDB with code 1HLU. 1HLU *Bos taurus* was considered to be the standard sequence for its higher sequence coverage and low E-value. The ClustalX programme provided an integrated environment for performing multiple sequence alignments and analyzing the results. All the 14 amino acid sequences (11 amino acid sequences from fishes & 3 from *Bos taurus*) of the β- actin protein, were loaded in Clustal X tool which provided a versatile colouring pattern for the display of the aligned sequences. The results of ClustalX programme was shown in Figure 1. However, the analysis of the result has shown a negligible amount of conservation in amino acid residues of the β- actin protein.

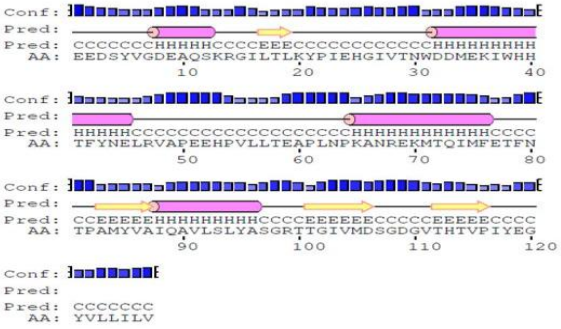

Fig. 4. 2D structure of β-actin Protein from *Chanos chanos* predicted by Psipred.

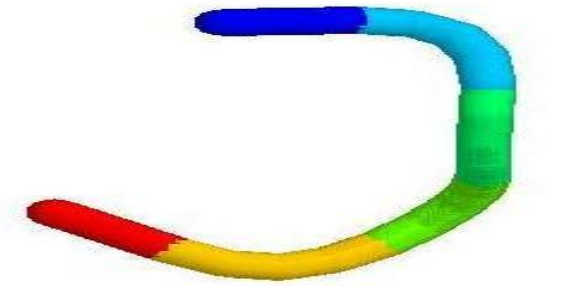

Fig. 5. Cartoon format of Model 1 (*Labeo rohita)* In Rasmol viewer.

The results obtained from multiple sequence analysis as shown in Figure 1, reflects only 2 conserved positions, in the form of ☐☐and :. Star represents completely conserved positions whereas, the colon represents very conservative substitutions. Of the 2 conserved residues, there was only 1 completely conserved position and 1 acceptable substitution was detected.

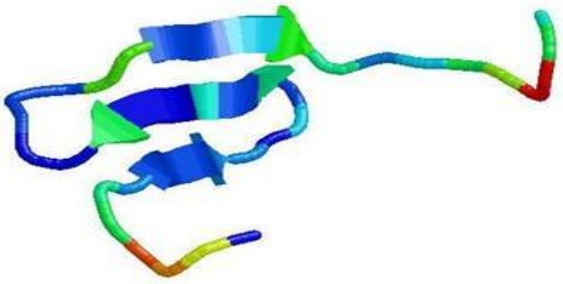

Fig. 6. Cartoon format of Model 2 (*Labeo rohita)* In Rasmol viewer.

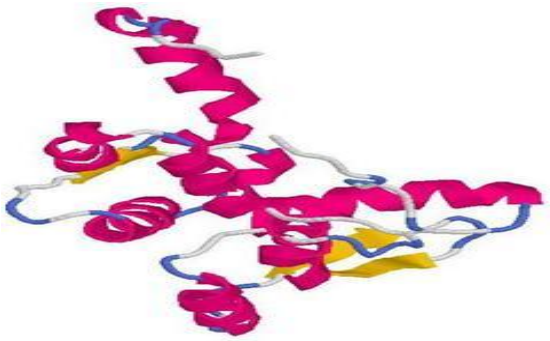

Fig. 7. Cartoon format of Model 3 (*Labeo rohita)* In Rasmol viewer.

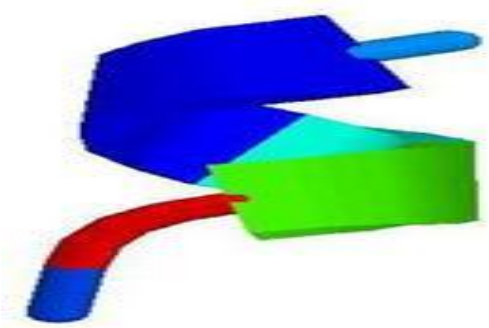

Fig. 8. Cartoon format (Shapely) of Model 4 (*Labeo rohita)* In Rasmol viewer.

The cladograms were constructed with the help of tree viewer tool (ClustalX) and Phylip. The trees that were constructed were unrooted trees. From the phylogram as presented in Figure 2**,** it is evident that 1HLU *Bos taurus* and 3 *Labeo rohita*, 2BTF *Bos taurus* and *Chanos chanos*, *1Labeo rohita*, 2 *Labeo rohita*, 4 *Labeo rohita* and *Catla catla* are located on a single branch. The phylogram of Phylip tool presented in Figure 3 demonstrates that 1HLU *Bos taurus*, 2OAN *Bos Taurus,* 2BTF *Bos taurus* and *Oreochromis mossambicus a*re located on a single branch. The above results conclusively emphasises the close phylogenetic relationship among these organisms with respect to amino acid sequence of β-actin protein. The secondary structure of all the 11 amino acid sequences of β-actin protein were predicted using tools like CFSSP, GOR, SOPMA, PSIPred listed in ExPASy (**E**xpert **P**rotein **A**nalysis **Sy**stem) proteomics server. The secondary structure tools have shown the alpha helices, beta pleated sheets and coils in all the 11 sequences at respective places. The number of secondary structures shown in a particular sequence using a spesific tool differed with respect to the other tool. In the present study, Table 1 indicates the 2D structures predicted by different secondary structure prediction tools. The 2D structure of the β-actin protein of one of the 11 sequences is presented in the form of Psipred as in Figure 4.

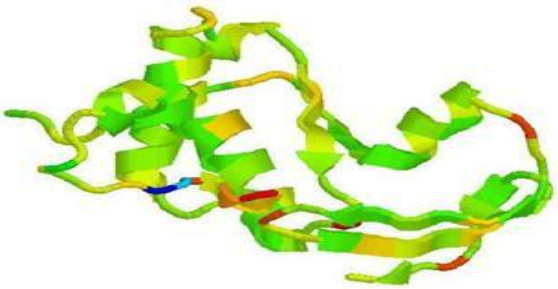

Fig. 9. Cartoon format (Temperature) of Model 5 (*Chanos Chanos*) In RasMol viewer.

Results of the present study using BLASTP analysis, demonstrates that 1HLU *Bos taurus* β-actin protein seems to bear closest resemblances with the β-actin protein of 8 fishes taken for study. The said sequence homology is possibly due to higher sequence similarity and sequence coverage with 11 protein sequences of the fishes.

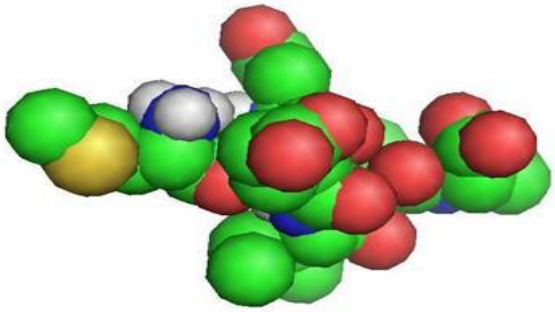

Fig. 10. Space fill format of Model 6 (*Catla catla*) In Pymol viewer.

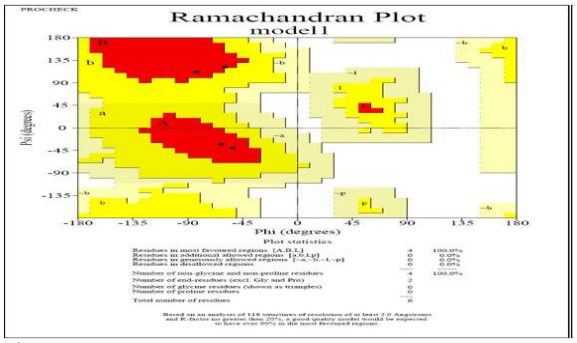

Fig. 11. Ramachandran Plot of Model1**(***Labeo rohita*).

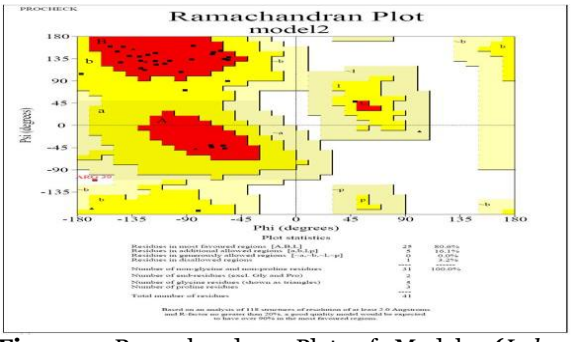

Fig. 12. Ramachandran Plot of Model2 **(***Labeo rohita*).

The results of multiple sequence alignment using ClustalX programme has depicted a negligible amount of conservation only at two sites like Methionine at position 285 and Isoleucine at 289. This negligible amount of conservation is probably be due to vast differences in length of the sequences and low sequence similarity among the 11 β-actin protein sequences of 8 fishes. Phylogenetic interrelationship of the fishes based on the β-actin protein, it has been observed that organisms of same taxon are laid on different branches, which is possibly because of least percentage of sequence sharing among the individuals of the same taxon.

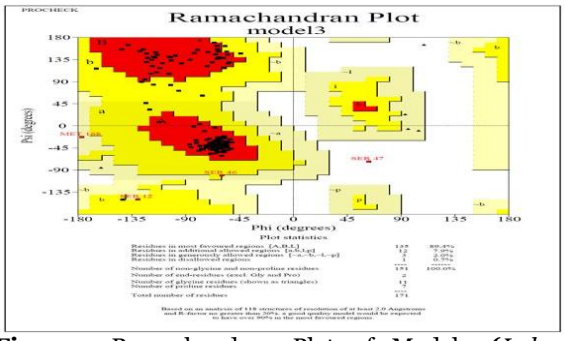

Fig. 13. Ramachandran Plot of Model 3**(***Labeo rohita*).

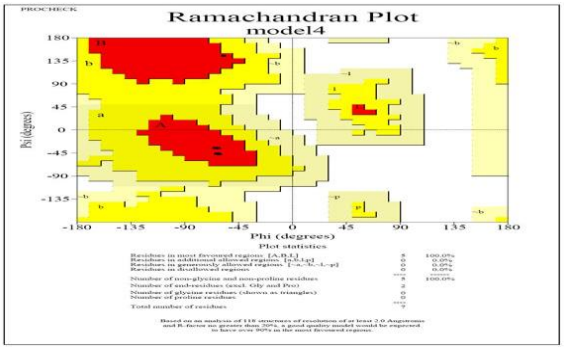

Fig. 14. Ramachandran Plot of Model 4 (*Labeo rohita*).

The secondary structures that were predicted using CFSSP, GOR and SOPMA were based on rule and propensity of amino acid residues in a particular secondary structure. All these tools have predicted structures which are more or less similar to each other. It is also revealed that the protein sequences of *Labeo calbasu, Heteropneustes fossilis, Clarias batracus, Lates calcarifer and Oreochromis mossambicus* have adequate variability in secondary structure compared to other sequences.

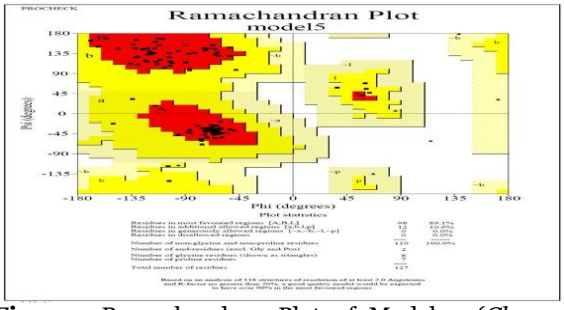

Fig. 15. Ramachandran Plot of Model 5 (*Chanos Chanos*).

The 3D structures of all the 11 β-actin protein sequences (Figure 5-10) have stated about 47-87% of structural similarity by SAVES server after energy minimization, which is confirmed by Ramchandran plot (Figure 11-15) that showed all the residues in the allowed regions.

Acknowledgements

The authors thankfully acknowledge DBT, Govt. of India for financial assistance and to the Head, P.G.Dept. of Zoology for providing necessary laboratory facilities.

References

Andreassen TK, Skjoedt K, Korsgaard B. 2005. Upregulation of estrogen receptor- α and vitellogenin in eelpout (*Zoarces viviparus*) by waterborne exposure to 4-tert-octylphenol and 17β-estradiol. Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology. **140,** 340-346. http://dx.doi.org/10.1016/j.cca.2005.03.003

Bernstein BW, Bamburg JR. 2010. ADF/cofilin: a functional node in cell biology. Trends in Cell Biology **20,** 187–195.

http://dx.doi.org/10.1016/j.tcb.2010.01.001

Cao J, Cao Z, Wu T. 2007. Generation of antibodies against DMRT1 and DMRT4 of *Oreochromis aurea* and analysis of their expression profile in *Oreochromis aurea* tissues. Journal of Genetics and Genomics **34,** 497-509.
http://dx.doi.org/10.1016/S1673-8527(07)60055-1

Condeelis J. 1993. Life at the leading edge: the formation of cell protrusions. Annual Reviews of Cell Biology **9,** 411-44.
http://dx.doi.org/10.1146/annurev.cb.09.110193.002211

Doi Y, Shinzawa N, Fukumoto S, Okano H, Kanuka H. 2010. ADF2 is required for transformation of the ookinete and sporozoite in malaria parasite development. Biochemical and Biophysical Research Communications. **397(4),** 668–672.
http://dx.doi.org/10.1016/j.bbrc.2010.05.155

Herman IM. 1993. Actin isoforms. Current Opinion in Cell Biology **5,** 48–55.
http://dx.doi.org/10.1016/S0955-0674(05)80007-9

Hofmann WA, Stojiljkovic L, Fuchsova B,. Vargas GM, Mavrommatis E, Philimonenko V, Kysela K, Goodrich JA, Lessard JL, J HT, Hozak P, Primal de Lanerolle. 2004. Actin is part of pre-initiation complexes and is necessary for transcription by RNA polymerase II. Nature Cell Biology **6,** 1094–1101.
http://dx.doi.org/10.1038/ncb1182

Kislauskis EH, Xiao-chun Zhu, Singer RH. 1997. β-Actin messenger RNA localization and protein synthesis augment cell motility. Journal of Cell Biology **136(6),** 1263–1270.
http://dx.doi.org/10.1083/jcb.136.6.1263

Lee JS. 2000. The internally self-fertilizing hermaphroditic teleost Rivulus marmoratus (Cyprinodontiformes, Rivulidae) β-actin gene: Amplification and sequence analysis with conserved primers. Marine Biotechnology **2,** 161–166.

Makioka A, Kumagai M, Hiranuka K, Kobayashi S, Takeuchi T. 2011. Entamoeba invadens: identification of ADF/cofilin and their expression analysis in relation to encystation and excystation. Experimental Parasitology **127(1),** 195–201.
http://dx.doi.org/10.1016/j.exppara.2010.07.018

Pollard TD, Borisy GG. 2003. Cellular motility driven by assembly and disassembly of actin filaments. Cell **112,** 453–465.
http://dx.doi.org/10.1016/S0092-8674(03)00120-X

Quitschke WW, Lin ZY, DePonti-Zilli L, Paterson BM. 1989. The β-actin promoter: high levels of transcription depend upon a CCAAT binding factor. Journal of Biological Chemistry **264(16),** 9539–9546.

Reece KS, McElroy D, Wu R. 1992. Function and evolution of actins. Evolutionary Biology **8(26),** 1-34.

Sohn, YC, Kobayashi M, Aida K. 2001. Regulation of gonadotropin β subunit gene expression by testosterone and gonadotropin-releasing hormones in the goldfish, Carassius auratus. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **129B,** 419–426.
http://dx.doi.org/10.1016/S1096-4959(01)00342-6

Sheterline P, Sparrow J. 1994. Actin. Protein Profile **1,** 1-121.

Stossel TP. 1994. The machinery of cell crawling. Scientific American. **271(3), 54-5,** 58-63.
http://dx.doi.org/10.1038/scientificamerican0994-54

Straub FB. 1942. Actin II Studies, Institute of Medical Chemistry University Szeged. **2**, 3-15.

Wang Y, Ge W. 2003. Gonadotropin regulation of follistatin expression in the cultured ovarian follicle cells of zebrafish, *Danio rerio*. General and Comparative Endocrinology **134,** 308–315. http://dx.doi.org/10.1016/S0016-6480(03)00275-2