

E-ISSN: 2320-7078 P-ISSN: 2349-6800 JEZS 2018; 6(4): 1444-1453 © 2018 JEZS Received: 15-05-2018 Accepted: 16-06-2018

Manchi Rajesh ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Biju Sam Kamalam

ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Alexander Ciji

ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Md. Shabaz Akhtar ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Debajit Sarma ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Narrotam Prasad Sahu

ICAR-Central Institute of Fisheries Education, Mumbai, Maharashtra, India

Atul Kumar Singh ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Correspondence Manchi Rajesh ICAR-Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand, India

Journal of Entomology and Zoology Studies

Available online at www.entomoljournal.com



Molecular cloning and *in silico* analysis of elongation factor 1A in *Schizothorax richardsonii*

Manchi Rajesh, Biju Sam Kamalam, Alexander Ciji, Md. Shabaz Akhtar, Debajit Sarma, Narrotam Prasad Sahu and Atul Kumar Singh

Abstract

In this manuscript, we report the cloning and *in silico* characterization of elongation factor 1A in a coldwater Trans-Himalayan cyprinid, *Schizothorax richardsonii*. The sequenced elongation factor 1A (*SrEF1A*) consisted of 1290 bp long partial ORF, which included a start codon (ATG) and 430 deduced amino acids. The *SrEF1A* sequence showed high homology (\geq 90%) with other teleosts and higher vertebrates. Phylogenetic analysis, multiple sequence alignment and prediction of conserved residues indicated a close evolutionary relationship among the cyprinids and conservation of *SrEF1A* protein across the vertebrate class. The deduced *SrEF1A* protein did not contain any signal peptide, but had two potential N-glycosylation motifs at 284th and 314th amino acid residue. Presence of many serine, threonine and tyrosine phosphorylation sites was also predicted, suggesting a potential post-translational regulation of the *SrEF1A* protein. *In silico* predictions of sub-cellular localization, function and protein-protein network illustrate the role of *SrEF1A* in the protein synthesis machinery of the cell. Finally, a reliable tertiary structure of *SrEF1A* protein was predicted with ten helixes and nineteen beta sheets. Ligand (GDP) binding sites in the tertiary structure were predicted at 15-22, 154, 156, 157 and 194-196 amino acid residues.

Keywords: Snow trout, elongation factor-1 alpha, house-keeping gene, reference gene, protein translation

1. Introduction

Elongation factor 1A (EF1A) is an important cellular protein involved in the decoding of mRNA codons to amino acids and thereby functional proteins ^[1]. It belongs to the family of Gprotein and is one of the four subunits found in eukaryotic elongation protein complex ^[2, 3] In fact, EFIA is one of the most abundant protein found in the cytoplasm of a normal growing cell, accounting for 1-3% of total protein [4, 5]. The GTP bound form of EF1A delivers aminoacyl-t RNA to the A site of ribosome, and after the correct codon and anticodon pairing of mRNA and tRNA, a conformational change in EF1A results in hydrolysis of GTP to GDP and release of EF1A-GDP complex from A site. The GDP of EF1A is actively exchanged for GTP by *EF1B* resulting in active *EF1A* which participate in another round of elongation [1, 6]. EFIA is attributed with many other functional roles apart from protein syntheses such as interaction with cytoskeleton proteins or actin binding property and the ability to bundle Factin [7], microtubule severing [8] and its stabilization [9]. The above functions suggest that *EF1A* plays a secondary role as a regulator of cytoskeletal rearrangements ^[8], but its biological significance is uncertain. Further, a large body of evidence supports the role of EF1A in nuclear export of t-RNA and shuttling of EF1A between nucleus and cytoplasm ^[1]. The EF1A is also known to direct misfolded proteins from ribosome to the proteasome ^[10]. Other functions include their role in apoptosis ^[11]; interaction with viral RNAs ^[12]; activator of protein kinases ^[13] and also altered expression of *EF1A* is implicated many kinds of tumors ^{[14,} ^{15]}. Other non-canonical functions of *EF1A* are detailed elsewhere ^[1]. These diverse functions suggest that EF1A is a multifunctional protein with many physiological roles.

EF1A has been characterised in many fishes (as listed in Table 1) and is highly conserved across the vertebrate class (Table1), which implies a corresponding functional conservation. A number of isoforms of *EF1A* are reported in plants and animals. For example, five *EF1A* isoforms have been discovered in Senegalese sole with different spatial expression pattern, though their functional significance is not yet known ^[16]. On the other hand, *EF1A* has been widely used as a reference gene or internal control for normalising the mRNA expression data

^[17, 18]. In the present study, we characterised 1290 bp of *EF1A* in *Schizothorax richardsonii*, in order to know the sequence of the transcript and protein, its phylogenetic relationship and protein structure. *S. richardsonii* is a coldwater cyprinid fish with wide distribution in the trans-Himalayan region. It forms an important component of subsistence fishery and is considered as a promising candidate species for aquaculture in Indian uplands.

2. Materials and methods

2.1. Amplification and cloning of *EF1A*

Total RNA was extracted from dorsal muscle tissue of *S. richardsonii* (weighing 15 g) using TRIzol reagent (Invitrogen, Carlsbad, CA, USA), as per manufacturers protocol. The extracted RNA was quantified using Take 3 system (Biotek Eon3 VT, USA). RNA quality was assessed in agarose gel electrophoresis using 1 μ g of total RNA. First strand cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems Thermo Fisher Scientific, MA, USA) and random hexamer, as per the manufacturer's instruction.

Gene-specific primers were designed based on the conserved regions identified by aligning available sequences of *EF1A* from closely related fishes of the cyprinid family, using Clustal omega online resource. Apt primers were chosen using Gene runner primer designing software (forward primer ATCAATCATGGGAAAGGAAAAG and reverse primer CCTCATGTCACGCACAGC). PCR was performed using the cDNA template, gene-specific primers, and Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, USA), as per manufacturers protocol. Amplified products were gel eluted, purified, ligated into pCR-4-TOPO vector and subsequently transformed into TOPO 10 cells (Invitrogen, Carlsbad, CA, USA). Positive colonies were screened and sequenced after plasmid extraction and confirmation through restriction digestion of plasmid.

2.2. In silico analysis of SrEF1A

The nucleotide sequence of SrEF1A was assembled in BioEdit version 7.0.5.2. The amino acid sequences were deduced using ExPASy (http://www.expasy.org/tools/). NetPhos 3.1 server was used to predict serine, threonine or tyrosine phosphorylation sites (http://www.cbs.dtu.dk/services/NetPhos/). N-glycosylation sites were predicted in the coding region using NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc server /NetNGlyc1.0Server/). CELLO2GO web server was used for protein subcellular localization prediction (http://cello.life.nctu.edu.tw/cello2go/). Protein-protein interaction was predicted using STRING web server (https://string-db.org/). Multiple sequence alignments were ClustalW2.0 performed using the program (http://www.ebi.ac.uk/tools/clustalw2) in CLC Genomics Workbench v.7.5.2 (CLC Bio, Aarhus, Denmark). Evolutionarily conserved residues of deduced protein were predicted using Consurf server (http://consurf.tau.ac.il/2016/). Homology modeling was performed using Swiss model ^[19] and I-TASSER online server [20]. I-TASSER COFACTOR was used to deduce ligand binding sites. The protein models were visualised in UCSF Chimera. The predicted structures validated **SAVES** were using v5.0 (http://servicesn.mbi.ucla.edu/SAVES/), Rampage Ramachandran plot (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) and ProSA-web server (https://prosa.services.came.sbg.ac.at/ prosa.php) for checking overall model quality (Z score) and local model quality (residue wise energy plot). Phylogenetic tree was constructed using the Neighbour-Joining method based on the deduced amino acid sequences of *SrEF1A* and other reported species *EF1A* in MEGA 7 software. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The evolutionary distances were computed using the Poisson correction method.

3. Results and Discussion

The PCR amplification resulted in an expected fragment of 1290 bp of EF1A (Fig.1), which was confirmed by sequencing. The nucleotide sequence of SrEF1A was deposited in NCBI Genbank under the accession number KU715836. The sequence result covered 5' start codon (ATG) but not 3' stop codon; thus the reported sequence is a partial cds. The deduced amino acid sequence of partial SrEF1A consisted of 430 amino acid residues. The full-length EF1A in other species is reported to be 462 in cyprinids, 461 in Xenopus, 461 in salmonids, 504 in elasmobranch, 462-463 in birds and 462 in mammals [16, 21] (number of residues as calculated from NCBI protein BLAST). The NCBI protein BLAST results further showed that >90% SrEF1A residues (Table 1) were similar across vertebrates, suggesting a resultant functional similarity. The phylogenetic tree constructed with 39 species EF1A including representatives of all vertebrate class resulted in three major clades (Fig. 2). The first clade included higher vertebrates, elasmobranchs and salmonids, whereas the second and third clade exclusively grouped the teleost EF1A. In the second clade, S. richardsonii EF1A was closely clustered with other cyprinids such as Ctenopharyngodon idella, Tor putitora, Cyprinus carpio, Carassius auratus and Danio rerio, suggesting a common evolutionary linkage. Multiple sequence alignment of SrEF1A with selected other vertebrate EF1A (Fig. 3) and the presence of conserved residues across a larger space (Fig. 4) suggests evolutionary conservation of EF1A protein in vertebrates.

The Cello2GO predicted subcellular localization (Fig. 5) indicates that SrEF1A protein is mainly localised in cytoplasm (68.8%). This is in agreement with its major role in protein synthesis which occurs in cytoplasm, as observed by other invitro studies ^[22]. Likewise, the functional prediction of SrEF1A protein (Fig. 6) indicates multiple functions which include translation factor, GTPase and hydrolase activities; and RNA, nucleic acid and ion binding. These predicted functions correlate exactly with their major role in protein synthesis. It is known that EF1A binds to t-RNA in its GTP form and then hydrolvsis of GTP ensures the fidelity of protein synthesis (GTP hydrolysis takes place only when correct pairing anticodon and codon is achieved); while Mg ion coordinates GDP+Pi state ^[23]. In any living system, protein-protein interaction is imperative to exert their proper function in vital processes such as cell growth, reproduction, motility, intracellular communication, and apoptosis ^[24]. As depicted in the interaction network diagram (Fig. 7), the major protein-protein interaction of SrEF1A included ribosomal protein L3, eukaryotic translation elongation factor 2b, eukaryotic translation elongation factor 1 gamma and other ribosomal protein, which are all involved in the protein synthesis machinery of eukaryotic cell ^[23, 6]. *SrEF1A* was also shown to interact with guanine nucleotide binding protein (G protein) and beta polypeptide 2-like 1 (GNB2L1), which

suggests their involvement in cellular processes other than protein synthesis such as signal transduction $^{[1, 25]}$. All the predicted protein-protein interactions were apparently significant with high confidence scores (> 0.9).

SrEF1A did not contain any signal peptides (Fig. 8), but the presence of two glycosylation sites at 284th (NITT motif) and 314th (NVSV motif) residue was predicted by NetNGlyc 1.0 server (Fig. 9). Generally, proteins without signal peptides are implausible to be exposed to N-glycosylation and thus SrEF1A may not be glycosylated (in vivo) despite the presence of potential motifs ^[26]. An in-depth in-vivo study is required to ascertain the possibility or lack of SrEF1A glycosylation. Another important post-translational mechanism that regulates protein activity occurs through phosphorylation and dephosphorylation, employing kinases and phosphatases ^[27]. Many studies have demonstrated that EF1A is subjected to phosphorylation by various kinases such as protein kinase C, multipotential S6 kinase, cyclin dependent kinase, casein kinase II and TGF-B receptor I which either up-regulates or down-regulates protein translation ^[28-30, 1]. In agreement, SrEF1A was predicted to contain many serine, threonine and tyrosine phosphorylation sites (Fig. 10), which suggests that SrEF1A activity might be regulated through post-translational phosphorylation.

The functional relevance of any protein at the molecular level is better understood through its three dimensional structure. When X-ray crystallography structure is unavailable, comparative homology modelling is commonly used to predict the structure of a query protein ^[20]. In this way, we also predicted the secondary and tertiary structure of *SrEF1A*. The predicted secondary structure of *SrEF1A* consisted 18.06 % helix, 35.81 % beta sheets and 45.58 % loop. The tertiary

structure of SrEF1A (amino acid residue 5-430; Fig 11A & B) was built based on comparative homology modelling using 5lzs mammalian ribosomal elongation complex with aminoacyl-tRNA, eEF1A and didemnin B as a template. The structure consisted of 10 helixes and 19 beta sheets. Ramachandran plot (Fig.12) of the predicted structure showed that 96.5 % of the amino acid residues were within the favourable region, 3.3% in the allowed region and only a single residue was outside the allowed region. Further, ProSA overall model quality (Z score = -8.86; Fig.13A) suggested that the predicted structure was in confirmation with the experimentally determined protein structure available in PDB. Similarly, local model quality (Fig. 13B) suggested that absence of problematic or erroneous regions in the predicted structure as there were no positive energy values (thick green line indicates residue wise energy plot calculated based on smoothened average energy score over each-40 residues). Taken together, all these structure validation parameters suggest that the predicted protein structure is reliable and satisfactory. To illustrate, superimposition of SrEF1A structure (Fig.11 C) with yeast elongation factor complex eEF1A (1f60) resulted in a RMSD value of 0.34 A° indicating high similarity between the structures. In addition, the ligand (GDP) binding sites in the SrEF1A protein structure were predicted as follows: HIS-15, VAL-16, ASP-17, SER-18, GLY-19, LYS-20, SER-21, THR-22, LYS-154, ASP-156, SER-157, SER-194, GLY-195, TRP-196 and they are clearly depicted in Figure 14A & B). Molecular simulation and dynamic studies are further required to understand the conformation changes that occur between GTP and GDP form of SrEF1A^[23].

Species	Similarity (%)	NCBI Accession
EF1A		
Schizothorax richardsonii	100	AQV11481
Tor putitora	99	AQV11477
Ctenopharyngodon idella	99	BAO56800
Cyprinus carpio	98	AAO49408
Danio rerio	97	NP_571338
Carassius auratus	97	BAB64567
Pimephales promelas	97	AAT91089
Clarias batrachus	97	BAO56800
Gobiocypris rarus	95	ADI60310
Oryzias latipes	93	XP_004074126
Salmo salar	92	NP_001117101
Oncorhynchus mykiss	91	NP_001117811
Fundulus heteroclitus	91	XP_012707957
Oreochromis niloticus	92	NP_001266576
Gadus morhua	90	ABD62881
Sparus aurata	91	AAD56406
Epinephelus coioides	91	AOW69105
Xenopus tropicalis	92	NP_989301
Gallus gallus	91	NP_001308445
Mus musculus	90	AAH04005
Bos taurus	90	CAA34756
Homo sapiens	90	BAD96702

Table 1: Similarity of SrEF1A with other teleost and higher vertebrate species

Journal of Entomology and Zoology Studies



Fig 1: Agarose gel image of amplified and cloned *SrEF1A* fragment. (A) Amplified PCR product of *SrEF1A*; (B) Cloned pCR 4.0 TOPO plasmid showing a clear insert (*SrEF1A*) of expected size (+1300 bp) after restriction digestion. M: marker, A: uncut intact plasmid with insert; B & C: restriction enzyme digested plasmid containing *SrEF1A* insert.



Fig 2: Phylogenetic tree constructed based on the deduced amino acid sequences of *SrEF1A* and other vertebrate *EF1A*, using MEGA 7 software. Numbers next to the branches indicate bootstrap values and scale bar represents evolutionary distance.

Journal of Entomology and Zoology Studies

S.richardsonii C.carpio D.rerio G.gallus H.sapiens B.taurus M.musculus	1 1 1 1 1	MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDIALWKF MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDI MGKEKTHINIVVIGHVDSGKSTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDI MGKEKTHINIVVIGHVDSGKSTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDI SLWKF MGKEKTHINIVVIGHVDSGKSTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDI SLWKF MGKEKTHINIVVIGHVDSGKSTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDISLWKF MGKEKTHINIVVIGHVDSGKSTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLKAERERGITIDISLWKF
S.richardsonii C.carpio D.rerio G.gallus H.sapiens B.taurus M.musculus	81 81 81 81 81 81	ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAGGVGEFEAGISKNGQTREHALLAFTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAGGVGEFEAGISKNGQTREHALLAFTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAGGVGEFEAGISKNGQTREHALLAFTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAGVGEFEAGISKNGQTREHALLAFTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGVKQLIVGVNKMDSTEP ETSKYYVTIIDAPGHRDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGVKQLIVGVNKMDSTEP
S.richardsonii C.carpio D.rerio G.gallus H.sapiens B.taurus M.musculus	161 161 161 161 161 161	PYSQARFEEITKEVSAYIKKIGYNPASVAFVPISGWHGDNMLEPSTNMGWFKGWKIERKEANANGUTLLDALDAILPPSR PYSQARFEEITKEVSAYIKKIGYNPASVAFVPISGWHGDNMLEPSTNMGWFKGWKIERKEGNANGUTLLDALDAILPPSR PYSQARFEEITKEVSAYIKKIGYNPASVAFVPISGWHGDNMLEASSNMGWFKGWKIERKEGNASGTLLDALDAILPPSR PYSQKYEEIVKEVSTYIKKIGYNPDTVFVPISGWNGDNMLEPSSNMPWFKGWKUTRKDGNSGTTLLDALDCILPPTR PYSQKYEEIVKEVSAYIKKIGYNPDTVAFVPISGWNGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEALDCILPPTR PYSQKYEEIVKEVSTYIKKIGYNPDTVAFVPISGWNGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEALDCILPPTR PYSQKRYEEIVKEVSTYIKKIGYNPDTVAFVPISGWNGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEALDCILPPTR PYSQKRYEEIVKEVSTYIKKIGYNPDTVAFVPISGWNGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEALDCILPPTR
S.richardsonii C.carpio D.rerio G.gallus H.sapiens B.taurus M.musculus	241 241 241 241 241 241 241 241	PTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPINTTTEVKSVEMHHESLTEATPGDNVGFNTKNVSVKDT PTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPINTTTEVKSVEMHHESLTBATPGDNVGFNVKNVSVKDT PTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPINVTTEVKSVEMHHESLSALPGDNVGFNVKNVSVKDV PTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPINVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDV PTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPINVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDV PTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPINVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDV
S.richardsonii C.carpio D.rerio G.gallus H.sapiens B.taurus M.musculus	321 321 321 321 321 321 321 321	RRGNVAGDSKNDPPQQAGSFNAQVIILNHPGQISQGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDNPKALKSGDAA RRGNVAGDSKNDPPKEAGSFNAQVIILNHPGQISQGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDNPKALKSGDAA RRGNVAGDSKNDPPKEAANFNAQVIILNHPGQISQGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDNPKALKSGDAA RRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAA RRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAA RRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAA RRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAA RRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAA
S.richardsonii C.carpio D.rerio G.gallus H.sapiens B.taurus M.musculus	401 401 401 401 401 401 401	IVEMEPGKPMCVESFSTYPPLGRFAVRDMR IVEMEPGKPMCVESFSTYPPLGRFAVRDMR IVEMVPGKPMCVESFSTYPPLGRFAVRDMR IVDMEPGKPMCVESFSDYPPLGRFAVRDMR IVDMVPGKPMCVESFSDYPPLGRFAVRDMR IVDMVPGKPMCVESFSDYPPLGRFAVRDMR

Fig 3: Multiple sequence alignment of *S. richardsonii EF1A* with other fishes and vertebrates. Protein sequence of *Cyprinus carpio* (AAO49408); *Danio rerio* (NP_571338); *Gallus gallus* (NP_001308445); *Homo sapiens* (BAD96702); *Bos tarus* (AAI05316); *Mus musculus* (AAH04005) were obtained from NCBI database.

1	11	21	31	41
GKEKIHINI	VVIGHVDSGK	STTTGHLIYK	CGGIDKRTIE	KFEKEAAEMG
eeeeeebbb	bbbbbbbeee	eebebebbbe	beebeeebe	ebeeebeeee
ff	ssssfff	ff s		
51	61	71	81	91
KGSEKYAWVL	DKLKAERERG	ITIDIALWKF	ETSKYYVTII	DAPGHRDFIK
eebbebbbbb	eeeeeeeee	ebbebbeeeb	eeeeebbbb	bbeeeebbe
fss	f fff ff	ssf	f s	ssfff f f
101	111	121	131	141
NMITGTSQAD	CAVIIIVAGGV	GEFEAGISKN	GOTREHALLA	FTLGVKQLIV
ebbebbbebe	bbbbbbbbee	eebeebbeee	eeeeebbbbb	eebebeebbb
fssfs fsf	S	ff	ff fs s	f fs f
151	161	171	181	191
GVNKMDSTEP	PYSQARFEEI	TKEVSAYIKK	IGYNPASVAF	VPISGWHCDN
bbeeeeeee	eeeeebeeb	eeebeebbee	eeeeeebeb	bbbbbbbbebee
ffff	f		f	s s f
201	211	221	231	241
MLEPSTNMG	FKGWKIERKE	ANANGVTLLD	ALDAILPPSR	PTDKPLRLPL
beeeeebeb	beebeeeeee	eeeeeebbe	bbeebeeeee	eeeebebeb
S	f	f		f
251	261	271	281	291
QDVYKIGGIG	TVPVGRVETG	VLKPGMVVTF	APANITTEVK	SVEMHHESLT
eeebeeeeb	ebeebebeeb	ebeeeebbb	beeeeeebe	ebeeeeebe
ff	s f s		f	
301	311	321	331	341
EATPEDNVGF	NIKNVSVKDI	RRGNVAGDSK	NDPPQQAGS	NAQVIILNHP
eeeeeebeb	ebeebeeeeb	eebebbbeee	eeeeebeeb	ebebbbbebe
fff		S	S	f
351	361	371	381	391
GQISQGYAPV	LDCHTAHIAC	KFAELKEKID	RRSGKKLEDN	PKALKSGDAA
bebeebebbb	bbbbebeeeb	ebeebeeebe	eeeeeeeee	eebbeeeeb
S	f			

401			41	1			421			
IVEM	[P <mark>G</mark> I	KPM	CV	SF	STY:	PP	LGR	TAV	RDMR	
bbebe	eee	eeb	bb	eebe	eeb	ee	bbel	obb	eeee	
			1	£			f		fff	
The co	nser	vatio	n sca	le:						
1	2	3	4	5	6	7	8	9		
Varia	ble		Ave	rage			Cons	erve	ed	

e - An exposed residue according to the neural-network algorithm.

b - A buried residue according to the neural-network algorithm.

f - A predicted functional residue (highly conserved and exposed).

 ${\bf s}$ - A predicted structural residue (highly conserved and buried).

Fig 4: Evolutionary conservation of SrEF1A amino acid residues as construed by ConSurf

Localization	Score
Extracellular	0.043
Plasmamembrane	0.065
Cytoplasmic	3.441
Cytoskeletal	0.010
ER	0.017
Golgi	0.013
Lysosomal	0.017
Mitochondrial	0.736
Chloroplast	0.174
Peroxisomal	0.109
Vacuole	0.022
Nuclear	0.352





Fig 6: Molecular function of SrEF1A as predicted by CELLO2GO



Predic	ted Functional partners:	Score
😁 rpl3	Ribosomal protein L3 (403 aa)	0.995
🔵 eef2b	Eukaryotic translation elongation factor 2b (858 aa)	0.995
🖲 eef1g	Eukaryotic translation elongation factor 1 gamma; Probably plays a role in anchoring the complex to other cellular components (486 aa)	0.994
🔵 rps2	Ribosomal protein S2 (280 aa)	0.987
🔵 rpl4	Ribosomal protein L4 (375 aa)	0.985
🔵 rpl7a	Ribosomal protein L7a (266 aa)	0.985
🔵 gnb2l1	Guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1; Involved in the recruitment, assembly and/or regulation of a variety of signalin	0.982
🔵 rpsa	Ribosomal protein SA; Required for the assembly and/or stability of the 40S ribosomal subunit. Required for the processing of the 20S rRNA- precursor	0.980
🔵 rplp0	Ribosomal protein, large, P0; Ribosomal protein P0 is the functional equivalent of E.coli protein L10 (316 aa)	0.970
🔵 rps3	Ribosomal protein S3 (245 aa)	0.965

Fig 7: Interaction network of SrEF1A protein showing different functional partners as predicted by STRING (with confidence scores)



Fig 8: The deduced signal peptide of SrEF1A protein (SignalP-4.1)



Fig 9: Potential N-glycosylation sites found in the SrEF1A protein



Fig 10: Potential serine, threonine, and tyrosine phosphorlyation sites found in the SrEF1A protein



Fig 11: Predicted tertiary structure of *SrEF1A*. (A) front view; (B) back view; (C) superimposition of model *SrEF1A* structure with yeast elongation factor complex *EF1A* (1f60).



Fig 12: Ramachandran plot (Rampage) showing dihedral angles Psi and Phi of amino acid residues in the predicted tertiary structure of SrEF1A.



Fig 13: ProSA overall model quality/Z score plot (A) and ProSA local model quality/residue-wise energy plot (B) of the predicted tertiary structure of *SrEF1A*.



Fig 14: Ligand binding sites in the predicted SrEF1A tertiary structure (A); closer view of the ligand binding sites (B).

4. Conclusion

In conclusion, we characterized 1290 bp long coding region of Elongation factor 1A (SrEF1A) in a Himalayan cyprinid, Schizothorax richardsonii. Phylogenetic analysis showed that SrEF1A clustered along with other cyprinids, indicating their common ancestral root. Moreover, SrEF1A sequence showed high similarity (> 90%) with that of other vertebrates, suggesting its evolutionary conservation. The presence of potential glycosylation and phosphorylation sites in SrEF1A indicates the possibility of post-translational activation of the protein through phosphorylation. The predicted sub-cellular localization, function and protein-protein network coherently emphasizes the role of SrEF1A in the protein translation machinery of the cell. Besides, we also predicted a reliable tertiary structure of the SrEF1A protein along with its ligand binding sites. Further studies on full-length cloning, molecular docking, simulation and dynamics of SrEF1A will reveal more subtle information about this protein and help us get a deeper insight into its molecular function.

5. Acknowledgements

R.M. gratefully acknowledges the Asian Fisheries Society -Kanazawa Ph.D. research grant (Asian Fisheries Society, Malaysia) for supporting his doctoral work. We are also thankful to the Indian Council of Agricultural Research for funding the study under the institutional research programme (AQ-17).

6. References

- 1. Sasikumar AN, Perez WB, Kinzy TG. The many roles of the eukaryotic elongation factor 1 complex. Wiley Interdisciplinary Reviews: RNA. 2012; 3(4):543-555.
- Browne GJ, Proud CG. Regulation of peptide-chain elongation in mammalian cells. European Journal of Biochemistry. 2002; 269:5360-5368.
- 3. Ejiri S. Moonlighting functions of polypeptide elongation factor1: from actin bundling to zinc finger protein R1-associated nuclear localization. Bioscience, Biotechnology, and Biochemistry. 2002; 66:1-21.
- 4. Slobin LI. The role of eukaryotic factor Tu in protein synthesis. The measurement of the elongation factor Tu content of rabbit reticulocytes and other mammalian cells by a sensitive radioimmunoassay. European Journal of Biochemistry. 1980; 110:555-563.
- Condeelis J. Elongation factor 1 alpha, translation and the cytoskeleton. Trends in Biochemical Sciences. 1995; 20:169-170.
- Berg J, Tymoczko TJ, Stryer L. Biochemistry. Edn 5. W.H. Freeman and Company, Madison Avenue, New York. 2003; 813-839.
- Yang F, Demma M, Warren V, Dharmawardhane S, Condeelis J. Identification of an actin-binding protein from Dictyostelium as elongation factor 1a. Nature. 1990; 347:494-496.
- 8. Shiina N, Gotoh Y, Kubomura N, Iwamatsu A, Nishida

E. Microtubule severing by elongation factor 1 alpha. Science. 1994; 266:282-285.

- 9. Moore RC, Durso NA, Cyr RJ. Elongation factor-1alpha stabilizes microtubules in a calcium/calmodulindependent manner. Cell Motility and the Cytoskeleton. 1998; 41(2):168-80.
- Gonen H, Dickman D, Schwartz AL, Ciechanover A. Protein synthesis elongation factor EF-1Alpha is an isopeptidase essential for ubiquitin-dependent degradation of certain proteolytic substrates. Advances in Experimental Medicine and Biology. 1996; 389:209-219.
- 11. Duttaroy A, Bourbeau D, Wang XL, Wang E. Apoptosis rate can be accelerated or decelerated by overexpression or reduction of the level of elongation factor-1 alpha. Experimental Cell Research. 1998; 238:168-176.
- Li Z, Pogany J, Tupman S, Esposito AM, Kinzy TG, Nagy PD. Translation elongation factor 1A facilitates the assembly of the tombus virus replicase and stimulates minus-strand synthesis. PLOS Pathogens. 2010; 6:e1001175.
- Yang W, Boss WF. Regulation of phosphatidylinositol 4kinase by the protein activator PIK-A49. Activation requires phosphorylation of PIK-A49. The Journal of Biological Chemistry. 1994; 269:3852-3857.
- 14. Lee JM. The role of protein elongation factor eEF1A2 in ovarian cancer. Reproductive Biology and Endocrinology. 2003; 1:69.
- 15. Thornton S, Anand N, Purcell D, Lee J. Not just for housekeeping: protein initiation and elongation factors in cell growth and tumorigenesis. Journal of Molecular Medicine. 2003; 81:536-548.
- 16. Infante C, Asensio E, Cañavate JP, Manchado M. Molecular characterization and expression analysis of five different elongation factor 1 alpha genes in the flatfish Senegalese sole (*Solea senegalensis* Kaup): Differential gene expression and thyroid hormones dependence during metamorphosis. BMC Molecular Biology. 2008; 9:19.
- Olsvik PL, Lie KK, Jordal AO, Nilsen TO, Hordvik I. Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. BMC Molecular Biology. 2005; 6:21.
- Infante C, Matsuoka MP, Asensio E, Cañavate JP, Reith M, Manchado M. Selection of housekeeping genes for gene expression studies in larvae from flatfish using realtime PCR. BMC Molecular Biology 2008, 9:28.
- Bordoli L, Schwede T. Automated protein structure modeling with Swiss-Model workspace and the protein model portal. Methods in Molecular Biology. 2012; 857:107-136.
- 20. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008; 9:40.
- ^{21.} Gao D, Li Z, Murphy T, Sauerbier W. Structure and transcription of the gene for translation elongation factor 1subunit alpha of zebrafish *Danio rerio*. Biochimica et Biophysica Acta. 1997, 1-5.
- ^{22.} Bohnsack MT, Regener K, Schwappach B, Saffrich R, Paraskeva E, Hartmann E *et al.* Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. The EMBO Journal. 2002; 21:6205-6215.
- 23. Shao S, Murray J, Brown A, Taunton J, Ramakrishnan V, Hegde RS. Decoding mammalian ribosome-mRNA states by translational GTPase complexes. Cell. 2017;

167:1229-1240.

- 24. Nooren IM, Thornton JM. Diversity of protein–protein interactions. The EMBO Journal. 2003; 22(14):3486-3492.
- 25. Mourtada-Maarabouni M. GNB2L1 (guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1). Atlas of Genetics and Cytogenetics in Oncology and Haematology. 2008; 12(2).
- Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT *et al.* Precision mapping of the human O-GalNAc glycoproteome through Simple Cell technology. The EMBO Journal. 2013; 32(10):1478-88.
- 27. Ardito F, Giuliani M, Perrone D, Troiano G, Lo Muzio L. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy. International Journal of Molecular Medicine. 2017; 40(2):271-280.
- 28. Peters HI, Chang YW, Traugh JA. Phosphorylation of elongation factor 1 (EF-1) by protein kinase C stimulates GDP/GTP-exchange activity. European Journal of Biochemistry. 1995; 234:550-556.
- 29. Venema RC, Peters HI, Traugh JA. Phosphorylation of valyl-tRNA synthetase and elongation factor 1 in response to phorbol esters is associated with stimulation of both activities. The Journal of Biological Chemistry. 1991; 266:11993-11998.
- Chang YW, Traugh JA. Insulin stimulation of phosphorylation of elongation factor 1 (eEF-1) enhances elongation activity. European Journal of Biochemistry. 1998; 251:201-207.