Indian Journal of Marine Sciences
Vol. 33(3), September 2004, pp. 231-237

Identification of the ribosomal proteins S20 and L10 from the amphioxus *Branchiostoma belcheri tsingtaunese* (Cephalochordata/Branchiostomidae)

Mei Liu[1, 2], Shicui Zhang[1]*, Zhenhui Liu[1] & Anlong Xu[3]

[1]Department of Marine Biology, Ocean University of China, Qingdao, 266003, P.R. China

[2]College of Life Sciences, Nanjing Normal University, Nanjing, 210097, P.R.China

[3]College of Life Science, Zhongshan University, Guangzhou, 510275, P.R. China

*[E-mail: zscshab@public.qd.sd.cn]

Received 29 October 2003, revised 5 July 2004

The complete cDNA and deduced amino acid sequences of the ribosomal protein S20 (*AmphiS20*) and L10 (*AmphiL10*) from the amphioxus *Branchiostoma belcheri tsingtaunese* are presented here. The *AmphiS20* cDNA consists of 591 base pair (bp) and encodes a 121 amino acid protein with a calculated molecular mass of 13,621 Da. The putative protein shares 65.6%-86.6% identity to the known eukaryotic homologues including animals, plants and fungi. The *AmphiL10* cDNA is 763 bp in length and encodes a 217 amino acid protein with a calculated molecular mass of 24,751Da. The deduced protein displays more than 65.4% sequence identity to its homologues examined. The high idenitity of AmphiS20 and AmphiL10 to their homologues from evolutionarily diverse organisms points to the remarkably conserved role these proteins play in ribosome structure and function. As house-keeping genes, determination of *AmphiS20* and *AmphiL10* will provide valuable normalizing tools for the study of transcriptional expression of other genes in amphioxus.

[**Key words**: Amphioxus, *Branchiostoma belcheri tsingtaunese*, ribosomal protein, S20, L10]
[**IPC code:** Int.Cl.[7] C12N 15/12]

Introduction

Ribosomes are the RNA-protein complexes that catalyze mRNA-directed protein synthesis in organisms. The structure of the ribosome is conserved throughout the prokaryotic and eukaryotic lineages, reflecting the early origin of their essential function. Each ribosome comprises two subunits. In eukaryotes, the large 60S subunit is composed of three ribosomal RNAs (rRNAs) and nearly 50 ribosomal proteins, whereas the small 40S subunit consists of one rRNA and about 30 proteins[1]. Most of the ribosomal proteins interact with multiple RNA elements to organize and stabilize the rRNA tertiary structures, adapting them for optimal function[2]. The primary structure of the ribosomal proteins has been determined in many organims. The comparative studies of the ribosomal proteins from different organims can reveal variations in potential sites for post-translational modification.

The genes encoding the ribosomal proteins S20 and L10 have been identified in eukaryotic species including animals, plants and fungi, respectively[3-6]. Amphioxus, a cephalochordate, is the closest extant relative to the vertebrates and has been widely known as the model animal to study the origin and evolution of the vertebrates[7]. To date, a large amount of genes including the ribosomal protein gene *AmphiS19* have been cloned and sequenced in amphioxus[8]. However, the genes coding for the ribosomal proteins L10 and S20 remains unknown. The present study reports the cloning and characterization of the ribosomal proteins S20 (*AmphiS20*) and L10 (*AmphiL10*) in the amphioxus *Branchiostoma belcheri tsingtauense*, which will provide valuable insight into the biochemical mechanism of protein synthesis in amphioxus and assist in our understanding of ribosomal function and evolution.

Materials and Methods

Adult amphioxus *Branchiostoma belcheri tsingtaunese* were collected from the sandy bottom of the sea near Shazikou, Qingdao, China, and starved for two days in sterilized filtered seawater to empty all the food in the gut. The gut was dissected out and frozen immediately in liquid nitrogen until use. The gut cDNA library was constructed according to the method of Liu *et al*[9]. Synthesized cDNA was ligated into pcDNA3-sfiI vector which had been modified

*For correspondence:
Tel.: +86-532-2032787 Fax: +86-532-2032787

from pcDNA3 vector (Invitrogen Inc.). About ~ 10^6 primary clones yielded approximately an equal number of colonies when transformed into *Escherichia coli* DH5α cells.

cDNA clones were randomly selected for sequencing. The insert length of each selected clone was examined by polymerase chain reaction (PCR) with the universal primers T7 (5′-TAATACGACT-CACTATAGGGA-3′) and SP6 (5′-ATTTAGGTG-ACACTATAGAA-3′) prior to plasmid DNA preparation. Both strands of all selected clones were sequenced with ABI PRISM 377XL DNA Sequencer and all sequences were then analyzed for coding probability with the DNATools program[10]. Multiple protein sequences were aligned using the MegAlign program by the CLUSTAL method in DNASTAR[11]. Accession numbers of the ribosomal protein sequences used for comparison were listed in Tables 1 and 2.

Results and Discussion

The first cDNA clone *AmphiS20* (accession number in Genbank: AF503586) consisted of 591 bp, which contained a 99 bp 5' untranslated region (UTR) in which one in-frame stop codon was located upstream of the first start codon ATG, a 126 bp 3' UTR and an open reading frame (ORF) of 366 bp (Fig. 1). The 5' UTR had an oligopyrimidine tract, which have been found to be present at the 5' end of many eukaryotic ribosomal protein mRNAs and may play a role in the regulation of their translation[12-14]. There was a polyadenylation signal AATAAA in the 3' UTR, which is required for the post-translational cleavage-polyadenylation of the 3' end of the pre-mRNA[15]. The open reading frame (ORF) encoded a putative 121 amino acid protein with a calculated molecular mass of 13,621 Da and an isoelectric point (pI) of 9.77. Totally, 22 out of the 121 amino acid residues were basic, and were scattered throughout the whole

Table 1—Representative members of the eukaryotic ribosomal protein S20 family

Protein	Organism (abbreviation)	Accession number	Amino acid	Source
S20Om	Rainbow trout, *Oncorhynchus mykiss* (Om)	AJ312336	119	EMBL
S20Hm	Human, *Homo sapiens* (Hs)	BC007507	119	GenBank
S20Rn	Norway rat, *Rattus norvegicus* (Rn)	X51537	119	EMBL
S20Xl	African clawed frog, *Xenopus laevis* (Xl)	M34706	119	GenBank
S20Ip	Channel catfish, *Ictalurus punctatus* (Ip)	AF402829	119	GenBank
S20Cf	Scallop, *Chlamys farreri* (Cf)	AF526250	119	GenBank
S20Dm	Fluit fly, *Drosophila melanogaster* (Dm)	Y11119	120	EMBL
S20At	Thale cress, *Arabidopsis thaliana* (At)	AY114024	119	GenBank
S20Sp	Fission yeast, *Schizosaccharomyces pombe* (Sp)	AL031798	119	EMBL
S20Bb	Amphioxus, *Branchiostoma belcheri* (Bb)	AF503586	121	GenBank

Table 2—Representative members of the eukaryotic ribosomal protein L10 family

Protein	Organism (abbreviation)	Accession number	Amino acid	Source
L10Dm	Fluit fly, *Drosophila melanogaster* (Dm)	O61231	218	Swissports
L10Bm	Wild silkworm, Bombyx mandarina (Bm)	AF099012	219	GenBank
L10Ip	Channel catfish, *Ictalurus punctatus* (Ip)	AF401563	215	GenBank
L10Pm	Sea lamprey, Petromyzon marinus (Pm)	AY130416	214	GenBank
L10Hs	Human, *Homo sapiens* (Hs)	P27635	214	Swissports
L10Mm	House mouse, *Mus musculus* (Mm)	P45634	214	Swissports
L10Bt	Cow, Bos taurus (Bt)	Q9XSI3	214	Swissports
L10Gg	Chicken, Gallus gallus (Gg)	Q08200	210	Swissports
L10Ce	Nematode, *Caenorhabditis elegans* (Ce)	Q09533	214	Swissports
L10At	Thale cress, Arabidopsis thaliana (At)	AY045866	220	GenBank
L10Sp	Fission yeast, *Schizosaccharomyces pombe* (Sp)	AL353014	221	EMBL
L10Bb	Amphioxus, *Branchiostoma belcheri* (Bb)	AY168758	216	GenBank

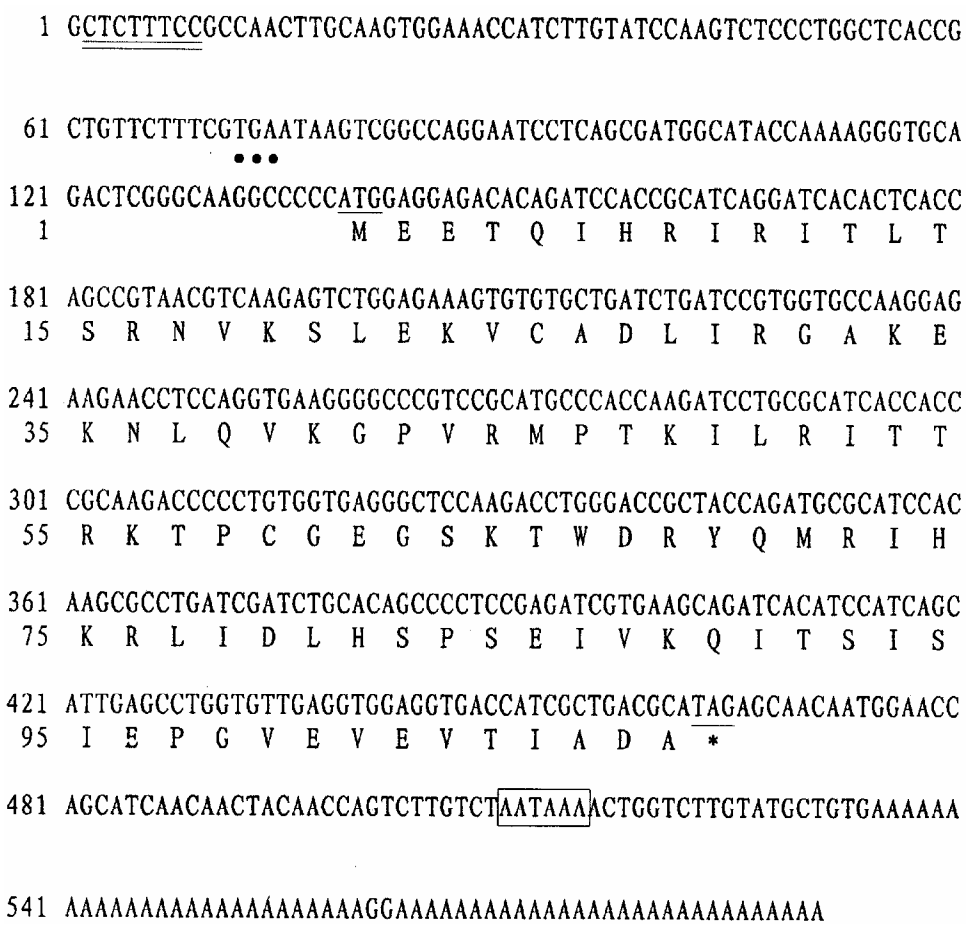

Fig. 1—Nucleotide and deduced amino acid sequence of *AmphiS20* (accession number in GenBank: AF503586). The presumed translational start and terminal sites are underlined, and asterisk represents the stop codon. The in-frame stop codon within 5' UTR is indicated as dots under the letters. The potential polyadenylation signal upstream the poly(A) tail is boxed and the oligopyrimidine tract within the 5'UTR is double underlined.

sequence of *AmphiS20*, which is a common feature for many ribosomal proteins[16]. These agree well with the earlier data from human, rat, *Xenopus*, rainbow trout, catfish, scallop, *Drosophila*, thale cress and yeast.

The deduced protein sequence of *AmphiS20* was compared with those of other known S20 from diverse organisms (Fig. 2). It showed that at the amino acid level, *AmphiS20* shared more than 86.6% identity to its homologues in the vertebrates such as human, rat, *Xenopus*, rainbow trout and catfish, more than 75.0% identity to those in the invertebrates including scallop and *Drosophila*, and more than 65.6% to those in other eukaryotes like thale cress and yeast. Of particular, the C-terminal 103 amino acids of *AmphiS20* had extremely high identity to those of other S20 ribosomal proteins with 64 being completely identical. This suggests an essential function for this portion of the protein.

The phosphorylation states of several ribosomal proteins have been shown to be important for their functions[17-21]. It has been found that *Drosophila* S20 has four potential protein kinase C phosphorylation sites (at residues 27-29, 62-64, 66-68 and 67-69, respectively), one potential protein kinase A phosphorylation site (at residues 101-104) and one potential casein kinase II phosphorylation site (at residues 95-98). All these phosphorylation states are also found in rainbow trout, human, rat, *Xenopus* and catfish S20 ribosomal proteins (Fig. 2). In contrast, there are only three potential protein kinase C phosphorylation sites (at residues 27-29, 66-68, 67-69), one potential protein kinase A phosphorylation site (at residues 101-104) and one potential casein kinase II phosphorylation site (at residues 95-98) in amphioxus, scallop, thale cress and yeast S20 ribosomal proteins. The significance for the presence or absence of the potential protein kinase C

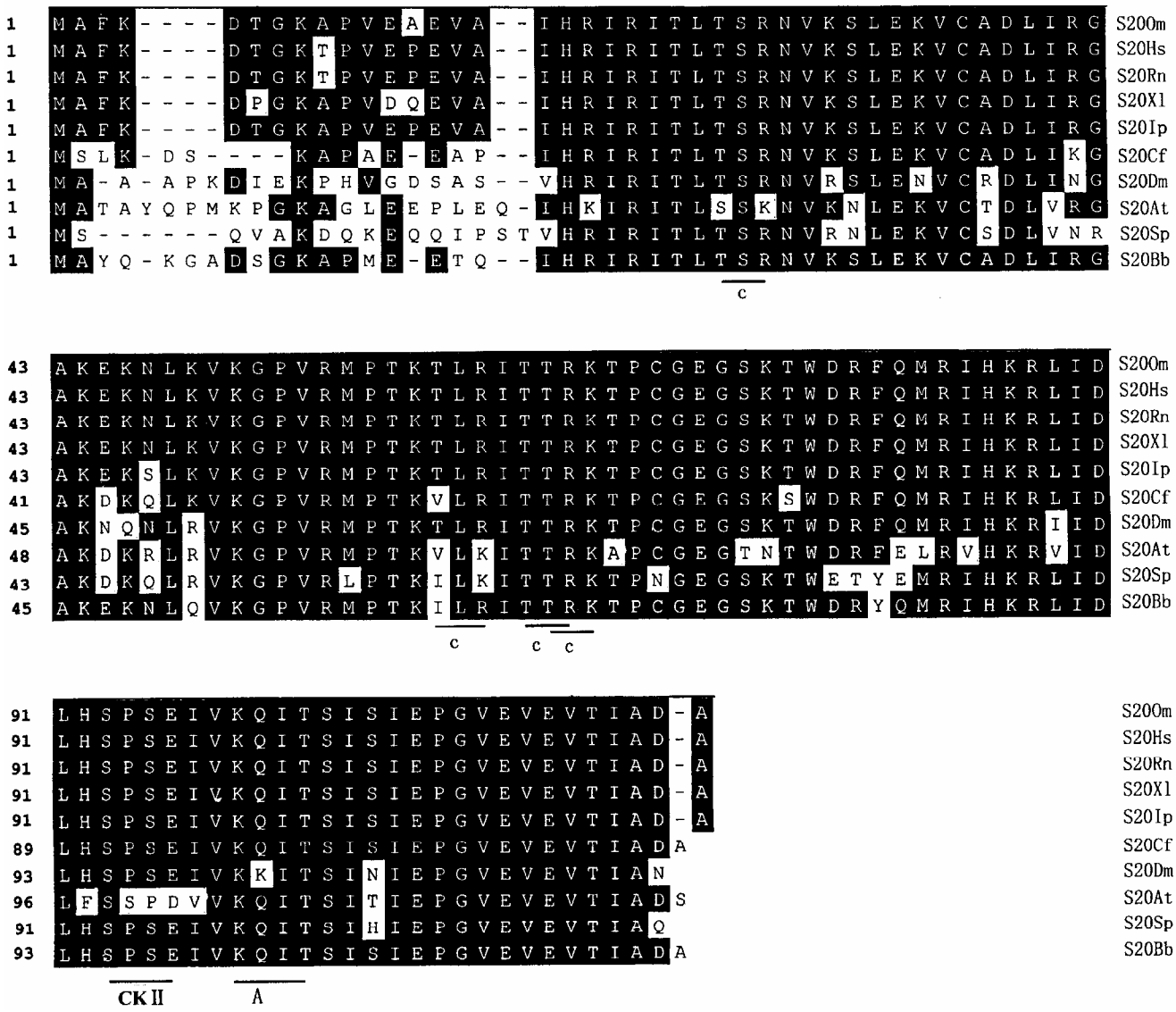

Fig. 2—Alignment of the amino acid sequences of the S20 from animals, plants and fungi by the method of CLUSTAL in DNAstar. Shaded (with solid black) residues are the amino acids that match the consensus. Gaps introduced into sequences to optimize alignments are represented by (-). Potential protein kinase C, protein kinase A and casein kinase II phosphorylation sites are underlined and designated below the residues as C, A, CKII respectively. See Table 1 for sequence reference.

phosphorylation site at residues 62-64 in different organisms remains a mystery and demands further study.

The second cDNA clone *AmphiL10* (accession number in Genbank: AY168758) was 763 bp long, including a 5' UTR of 17 bp, an ORF of 651 bp and a 3' UTR of 95 bp. An oligopyrimidine tract was located at nucleotides 2-8 and a putative polyadenylation signal at nucleotides 706-711 (Fig. 3). The putative *AmphiL10* protein consisted of 216 amino acids with a calculated molecular mass of 24,751 Da. It is rich in basic amino acids (44/216) and has an isoelectric point (pI) of 10.03. In addition, the tetrapeptide RXXR, which is critical for nucleolar localization[22], was found three times at residues 4-7, 7-10, 21-24, respectively, in *AmphiL10*. These observations agree with earlier reports from human, mouse, cow, chicken, catfish, lamprey, *Drosophila*, silkworm, nematode, thale cress and yeast.

Comparison of the deduced *AmphiL10* amino acid sequence with other known L10 protein sequences revealed greater than 65.4% sequence identity between species including human, mouse, cow, chicken, catfish, lamprey, *Drosophila,* silkworm,

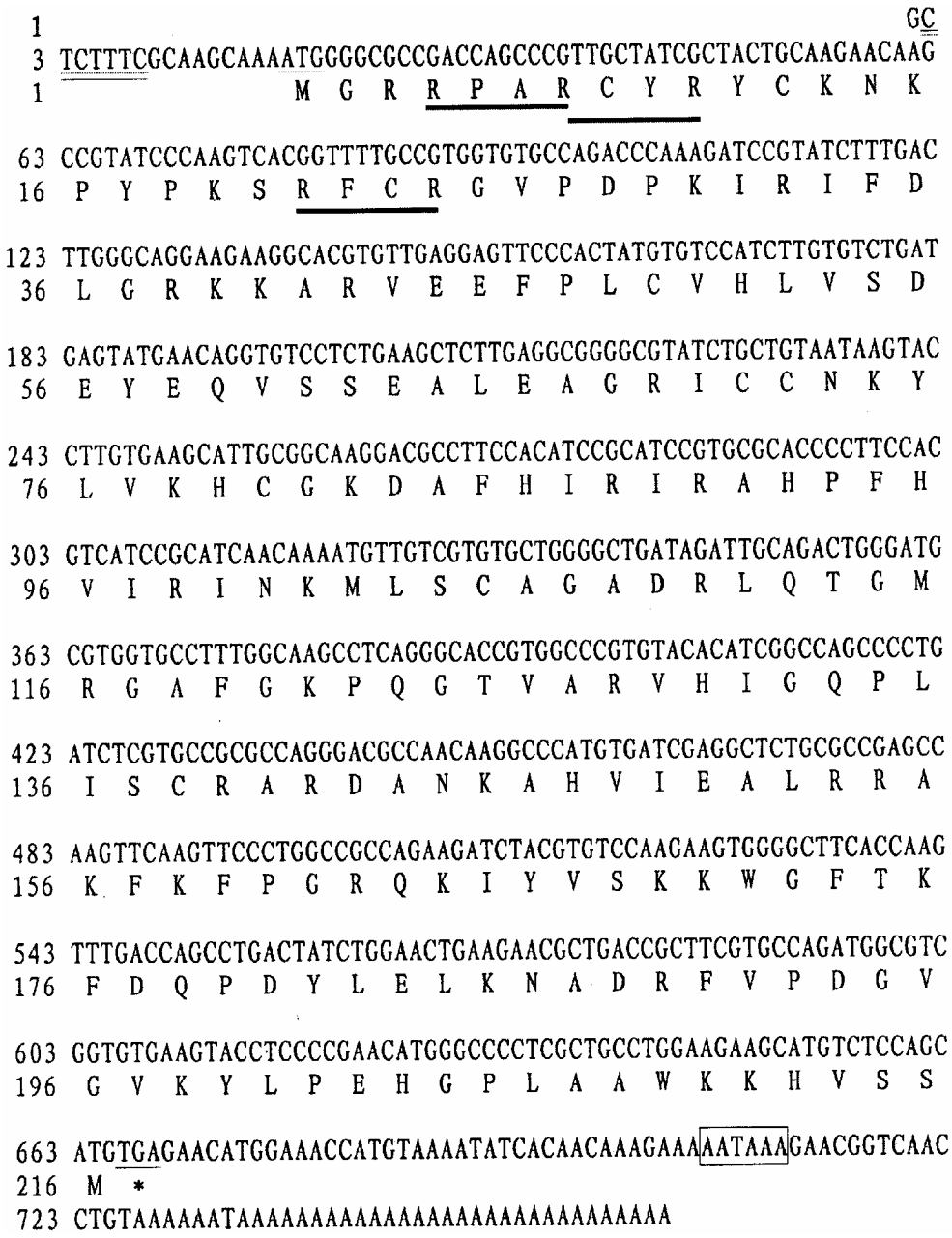

Fig. 3—Nucleotide and deduced amino acid sequence of *AmphiL10* (accession number in GenBank: AY168758). The presumed translational start and terminal sites are underlined, and asterisk represents the stop codon. The potential polyadenylation signal upstream the poly(A) tail is boxed and the oligopyrimidine tract within the 5'UTR is double underlined. The putative nucleolar localization signal (RXXR) in the AmphiL10 is marked by a heavy bar.

nematode, thale cress and yeast. As noted earlier, the highest level of conservation was concentrated in the N-terminal 175 residues with a total of 97 amino acids demonstrating complete conservation between different species. Two potential protein kinase C phosphorylation sites were identified at residues 137-139 and 168-170, and one potential casein kinase II phosphorylation site at residues 174-177 (Fig. 4). The potential protein kinase C phosphorylation sites are conserved in all the species examined.

Ribosomal protein L10 has been shown to be peripherally associated with the 60S subunit[23], which is essential for the joining of the 40S and 60S subunits to form a functional 80S ribosome[24]. Besides, the deduced amino acid sequence of *AmphiL10* was homologous to QM proteins from human, rat and

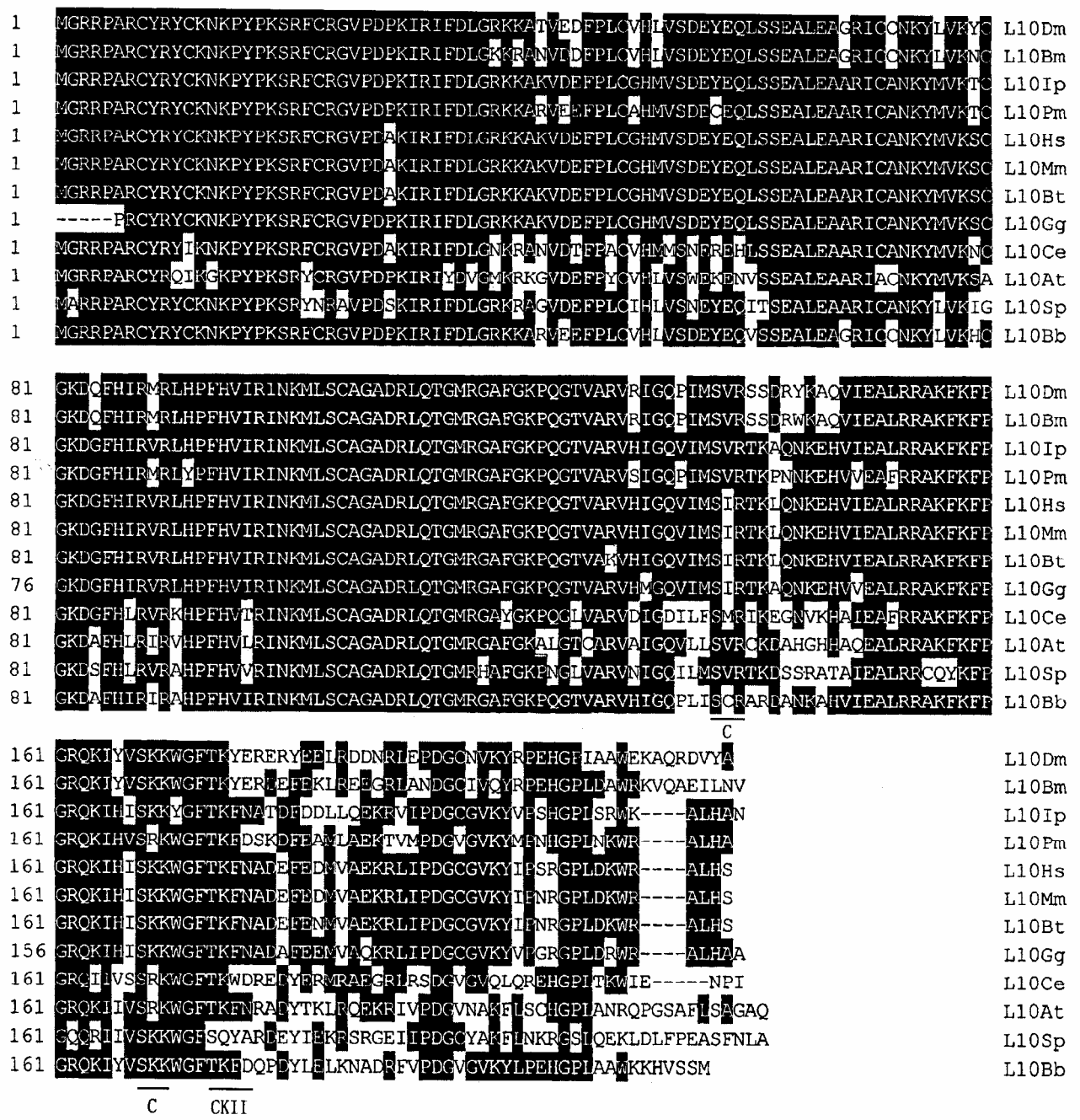

Fig. 4—Alignment of the amino acid sequences of the L10 from animals, plants and fungi by the method of CLUSTAL in DNAstar. Shaded (with solid black) residues are the amino acids that match the consensus. Gaps introduced into sequences to optimize alignments are represented by (-). Potential protein kinase C and casein kinase II phosphorylation sites are underlined and designated below the residues as C and CKII. See Table 2 for sequence reference.

Drosophila as revealed by BLASTX search. QM protein plays a role in both control of cell growth and proliferation, perhaps as a tumor suppressor, and in energy metalolism[25-27]. It is therefore possible that in addition to its function in the joining of the 40S and 60S subunits, L10 including *AmphiL10* has an extra-ribosomal function in organisms.

In this study, both *AmphiS20* and *AmphiL10* showed high identity to their known counterparts from diverse organisms, suggesting a crucial role they play in ribosomal structure and function. As constitutively expressed genes, ribosomal proteins have often been used to normalize gene expression[28]. Identification of *AmphiS20* and *AmphiL10* may

provide useful normalizing tools to study the transcriptional expression of other genes from amphioxus.

Acknowledgement

We thank Dr. Jianwen Wei and Chunyan Mou in Life Science College of Zhongshan University for help in cDNA library construction. The work was supported by Ministry of Science and Technology (MOST) of China (2002AA629190).

References

1 Kay, M.A. & Jacobs-Lorena, M., Developmental genetics of ribosome synthesis in *Drosophila*, *Trends Genet*., 3 (1987) 347-351.

2 Kurland, C.G., Jorgensen, F., Richter, A., Ehrenberg, M., Bilgin, N. & Rojas, A.M., Through the accuracy window. In: *The ribosome: Structure, function and evolution*. edited by Hill, W.E. *et al*. (American Society for Microbiology, Washington, DC) 1990, pp. 513-526.

3 Chan, Y.L. & Wool, I.G., The primary structure of rat ribosomal protein S20, *Biochim. Biophys. Acta*, 1049 (1990) 93-95.

4 Patterson, A., Karsi, A., Feng, J. & Liu, Z., Translational machinery of channel catfish: II. Complementary DNA and expression of the complete set of 47 60S ribosomal proteins, *Gene*, 305 (2003) 151-160.

5 Nishi, R., Hashimoto, H., Uchimiya, H. & Kato, A., The primary structure of two proteins from the small ribosomal subunit of rice, *Biochim. Biophys. Acta*, 1216 (1993) 113-114.

6 Hermann-Le Denmat, S., Sipiczki, M. & Thuriaux, P., Suppression of yeast RNA polymerase III mutations by the URP2 gene encoding a protein homologous to the mammalian ribosomal protein S20, *J. Mol. Biol.*, 240 (1994) 1-7.

7 Stokes, M.D. & Holland, N.D., The lancelet: also known as "amphioxus," this curious creature has returned to the limelight as a player in the phylogenetic history of the vertebrates, *Am. Sci*, 86(1998) 552-560.

8 Liu, M., Zhang, S., Liu, Z., Yuan, J. & Xu, A. Molecular cloning and phylogenetic analysis of the ribosomal protein S19 from amphioxus *Branchiostoma belcheri tsingtaunese*, *Curr. Sci.*, 83 (2002) 1249-1253.

9 Liu, Z., Zhang, S., Yuan, J., Sawant, M.S., Wei. J. & Xu, A., Molecular cloning and phylogenetic analysis of AmphiUbf80, a new member of ubiquitin family from the amphioxus *Branchiostoma belcheri tsingtauense, Curr. Sci.,* 83 (2002) 101-104.

10 Rehm, B.H., Bioinformatic tools for DNA/protein sequence analysis, functional assignment of genes and protein classification, *Appl. Microbiol. Biotechnol.*, 57 (2001) 579-592.

11 Burland, T.G., DNASTAR's Lasergene sequence analysis software, *Methods Mol. Biol.*, 132 (2000) 71-91.

12 Mariottini, P., Bagni, C., Annesi, F. & Amaldi, F., Isolation and nucleotide sequences of cDNAs for *Xenopus laevis* ribosomal protein S8: similarities in the 5' and 3' untranslated regions of mRNAs for various r-proteins, *Gene*, 67 (1988) 69-74.

13 Colombo, P. & Fried, M., Functional elements of the ribosomal proteinL7a (rpL7a) gene promoter region and their conservation between mammals and birds, *Nucleic Acids Res*., 20 (1992) 3367-3373.

14 Avni, D., Shama, S., Loreni, F. & Mcyuhas, O., Vertebrate mRNAs with a 5'-terminal pyrimidine tract are candidates for translational repression in quiescent cells: Characterization of the translational *cis*-regulatory element, *Mol. Cell. Biol.*, 14 (1994) 3822-3833.

15 Proudfoot, N.J. & Brownlee, G.G., 3' Non-coding region sequences in eukaryotic messenger RNA, *Nature*, 263 (1976) 211-214.

16 Mariottini, P., Bagni, C., Annesi, F. & Amaldi, F., Isolation and nucleotide sequences of cDNAs for *Xenopus laevis* ribosomal protein S8: similarities in the 5' and 3' untranslated regions of mRNAs for various r-proteins, *Gene*, 67(1988) 69-74.

17 Banham, A.H., leader, D.P. & Smith, G.L., Phosphorylation of ribosomal proteins by the Vaccinia virus BIR protein kinase, *FEBS Lett*, 321(1993) 27-31.

18 Simonin, D., Diaz, J.J., Kindbeiter, K., Denoroy, L. & Madjar, J.J., Phosphorylation of ribosomal protein L30 after Herpes Simple Virus type I infection, *Electrophoresis*, 16(1995) 854-859.

19 Kim, J.M., Ch, J.Y., Marshak, D.R. & Bae, Y.S., Interaction of the beta subunit of casein kinase II with the ribosomal protein L5, *Bioch. Biophys. Res. Commun.*, 226(1996) 180-186.

20 Patel, H.R., Terada, N. & Gelfand, E.W., Rapamycin sensitive phosphorylation of ribosomal protein S17 by P70 S6 kinase, *Biochem. Biophys. Res. Commun.*, 227 (1996) 507-512.

21 Proud, C.G., P70 S6 kinase-an enigma with variations, *Trends Biochem. Sci.*, 21(1996) 181-185.

22 Quaye, I.K., Toku, S. & Tanaka, T., Sequence requirement for nucleolar localization of the rat ribosomal protein L31, *European J. Cell. Biol.*, 69 (1996) 151-155.

23 Dick, F.A., Karamanou, S. & Trumpower, B.L., QSR1, an essential yeast gene with a genetic relationship to a subunit of the mitochondrial cytochrome bc1 complex, codes for a 60S ribosomal subunit protein, *J. Biol. Chem.*, 272 (1997) 13372-13379.

24 Eisinger, D.P., Dick, F.A. & Trumpower, B.L., Qsrlp, a 60S ribosomal subunit protein, is required for joining of 40S and 60S subunits, *Mol. Cell. Biol.*, 17 (1997) 5136-5145.

25 Dowdy, S.F., Lai, K.M., Weissman, B.E., Matsui, Y., Hogan, B.L.M. & Stanbridge, E.J., The isolation and characterization of a novel cDNA demonstrating an altered mRNA level in nontumorigenic Wilms' microcell hybrid cells, *Nucl. Acids Res.*, 19 (1991) 5763-5769.

26 Koller, H.T., Klade, T., Murach, K.F. & Breitenbach, M., The yeast temperature-sensitive mutant grc5 is involved in metabolic control of growth and survival under starvation conditions, *Biol. Hoppe-Seyler*, 374 (1993) 772.

27 Koller, H.T., Klade, T., Ellinger, A. & Breitenbach, M., The yeast growth control gene *GRC5* is highly homologous to the mammalian putative tumor suppressor gene QM, *Yeast*, 12 (1996) 53-65.

28 Moshe, T., Michael, R.W., Detlef, B. & Chantal, D.V., Cloning and characterization of the L15 ribosomal protein gene homologue from the crayfish *Orconectes limosus*, *Biochem. Biophys. Res. Commun*., 256 (1999) 313–316.