Molecular cloning and phylogenetic analysis of the ribosomal protein S19 from amphioxus Branchiostoma belcheri tsingtauense

M. Liu, S. Zhang*, Z. Liu, J. Yuan and A. Xu†

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

An amphioxus cDNA, AmphSi19, encoding the ribosomal protein S19 was isolated from the gut cDNA library of Branchiostoma belcheri tsingtauense. The cDNA contains a 444 base pair (bp) open reading frame, flanked by a 27 bp 5' untranslated region and a 138 bp 3' untranslated region. The ORF encodes a putative 147 amino acid protein with a calculated molecular mass of 16,222 Da. Alignment of the complete amino acid sequences of 14 eukaryotic S19 revealed that AmphSi19 exhibited >71.0% similarity to all known vertebrate homologues and <59.9% to those of the other eukaryotes, including invertebrates. The phylogenetic analysis based on the amino acids of vertebrate and ribosomal S19 proteins showed that the amphioxus protein was at the base of a clade of vertebrate S19 proteins, indicating that amphioxus is not only the sister group of extant vertebrates but also the basal lineage of chordates.

RIBOSOMES are organelles that mediate the sequential addition of amino acids to the carboxyl end of the growing polypeptide chain, according to the blueprints encoded by the mRNA. Each ribosome consists of two subunits. The eukaryotic 80S ribosome is composed of a large subunit ~60S, and a small one ~40S, while the prokaryotic 70S ribosome has a large subunit ~50S, and a small one ~30S. The large subunit contains three ribosomal RNAs (rRNAs), 5S, 5.8S and 28S, in eukaryotes, but only two, 5S and 23S in prokaryotes. The small subunit contains a single RNA in both types of organism: an 18S rRNA in eukaryotes and a 16S rRNA in prokaryotes. Both eukaryotic and prokaryotic small subunits comprise several dozen ribosomal proteins. The proteins of the small subunit are called S1, S2... and those of the large subunit are called L1, L2... The ribosomal proteins have been largely identified and their sequences determined. Many ribosomal proteins have been shown to bind specific regions of rRNA. Ribosomal proteins catalyse ribosome assembly and stabilize rRNA tertiary structure, adapting the structure of the ribosome for optimal function. The sequences of most eukaryotic ribosomal proteins have counterparts in prokaryotic ribosomal proteins, suggesting that they might derive from common ancestral nucleotide sequences present before the divergence of eukaryotes and prokaryotes, and be well-conserved throughout evolution. However, it has been recently shown that the possibility of the phylogenetic utility of the ribosomal proteins cannot be ruled out.

The eukaryotic S19, a core protein that is associated with the 18S rRNA of the 40S small subunit generally contains 143 to 156 amino acids. There are no known counterparts in prokaryotes, mitochondria and chloroplast, and therefore this protein appears to be a recent addition to the eukaryotic ribosomal protein repertoire. The genes encoding eukaryotic ribosomal protein S19 have been identified extensively in species, including animals, plants, and fungi. Amphioxus, a cephalochordate, is the closest living relative to the vertebrate, and has been widely known as the most important animal to study the origin and evolution of vertebrates. To date, there are more than 100 species of amphibians and sequenced in amphioxus such as the Hox, Insulin-like (ILP), Cdx, and Amphitoxin. However, the *For correspondence. (e-mail: zscshab@public.qd.sd.cn)
Table 1. Representative members of the eukaryotic ribosomal protein S19 family

<table>
<thead>
<tr>
<th>Protein</th>
<th>Organism (abbreviation)</th>
<th>Accession number</th>
<th>Amino acid</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>S19Hs</td>
<td>Human, Homo sapiens (Hs)</td>
<td>P39019</td>
<td>144*</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Mm</td>
<td>House mouse, Mus musculus (Mm)</td>
<td>BC034506</td>
<td>145</td>
<td>GenBank</td>
</tr>
<tr>
<td>S19Rn</td>
<td>Norway rat, Rattus norvegicus (Rn)</td>
<td>P17074</td>
<td>144*</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19lp</td>
<td>Channel catfish, Ictalurus punctatus (lp)</td>
<td>AF403282</td>
<td>147</td>
<td>GenBank</td>
</tr>
<tr>
<td>S19Mg</td>
<td>Atlantic hagfish, Myxine glutinosa (Mg)</td>
<td>Q9Y9H3</td>
<td>145</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Bb</td>
<td>Amphioxus, Branchiostoma belcheri (Bb)</td>
<td>AF491451</td>
<td>147</td>
<td>GenBank</td>
</tr>
<tr>
<td>S19Dm</td>
<td>Flutty fly, Drosophila melanogaster (Dm)</td>
<td>P39018</td>
<td>155*</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Mu</td>
<td>Softshell, Mya arenaria (Mu)</td>
<td>Q94613</td>
<td>149</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Cc</td>
<td>Nematode, Caenorhabditis elegans (Ce)</td>
<td>O18650</td>
<td>146</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19As</td>
<td>Pig roundworm, Ascaris suum (As)</td>
<td>P24494</td>
<td>148</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Sp-A</td>
<td>Fission yeast, Schizosaccharomyces pombe (Sp)</td>
<td>P58234</td>
<td>144</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Sp-B</td>
<td>Fission yeast, S. pombe (Sp)</td>
<td>P79016</td>
<td>143</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Eh</td>
<td>Ameba, Entamoeba histolytica (Eh)</td>
<td>O15631</td>
<td>148</td>
<td>SWISS-PROT</td>
</tr>
<tr>
<td>S19Os</td>
<td>Rice, Oryza sativa (Os)</td>
<td>P40978</td>
<td>146</td>
<td>SWISS-PROT</td>
</tr>
</tbody>
</table>

*The initial methionine is removed.

Figure 1. Nucleotide and deduced amino acid sequence of AmphipS19 (accession number in GenBank: AF491451). The presumed translational start and terminal sites are underlined, and asterisk represents the stop codon. The potential polyadenylation signal upstream with respect to the poly(A) tail is boxed and the oligopyrimidine tract within the 5'UTR is double underlined. The basic amino acid cluster in AmphipS19 is marked by a heavy bar.

The nucleotide sequence was obtained from the clone 027 selected randomly. The initial BLAST search revealed that it was homologous to the genes coding for the ribosomal protein S19 from human and rat, and thus designed AmphipS19. Figure 1 shows the nucleotide and deduced amino acid sequence of the AmphipS19 cDNA (accession number in GenBank: AF491451). The cDNA consisted of 609 bp, which contained a 27 bp 5' untranslated region (UTR), a 138 bp 3' UTR and an open reading frame of 444 bp. The 5' UTR had an oligopyrimidine tract, which has been reported to be present at the 5' end of many eukaryotic ribosomal protein mRNAs and may play a role in the regulation of their translation. There was a polyadenylation signal AATAAA in the 3' UTR, which is required for post-translational cleavage-polyadenylation of the 3' end of the pre-mRNA.

The open reading frame encoded a putative 147 amino acid protein with calculated molecular mass of 16,222 Da. It is a basic protein with an isoelectric point of 9.1, which is similar to other basic proteins.
Figure 2. Alignment of amino acid sequences of S19 from animals, plants and fungi by CLUSTAL method in DNAsStar. Shaded (with solid black) residues are the amino acids that match the consensus. Gaps introduced into sequences to optimize alignments are represented by (–). See Table 1 for sequence reference.

(pl) of 10.44, a common feature of many ribosomal proteins. The basic and acidic amino acids were usually clustered together in ribosomal proteins; however, only four basic residues between positions 83 and 86 in AmphS19 are clustered together, and other basic and acidic amino acids are evenly distributed.

Fourteen full-length amino acid sequences of S19 from animals, plants and fungi were aligned using CLUSTAL
Figure 3. Tables of sequence similarity and divergence calculated with the CLUSTAL method in the DNASTar. See Table 1 for sequence reference.

Figure 4. Phylogenetic tree constructed from the amino acid sequences of the animal S19 by neighbour-joining. Branch length represents the evolutionary distance. The position of AmphiS19 is just intermediate between the vertebrate S19 and invertebrate S19, supporting the fact that amphioxus is an organism transitional from invertebrate to vertebrate. See Table 1 for sequence reference.

Method in the software package DNASTar. The accession numbers for these sequences are listed in Table 1. The alignment confirms that these proteins are highly conserved throughout the evolution of chordates, yet they are relatively divergent among eukaryotes (Figure 2). The tables of sequence similarity and divergence were also calculated with the CLUSTAL method (Figure 3). It showed that at the amino acid level, AmphiS19 exhibited >71.0% similarity to known vertebrate homologues and <59.9% to those of the other eukaryotes, including invertebrates. Apparently, AmphiS19 is more closely related to the vertebrate S19 rather than those in invertebrates.

The phylogenetic tree constructed from the complete amino acid sequences of AmphiS19 and the other ten animal S19s by neighbour-joining in DNASTar program demonstrated that the relationship among different species well reflected the established phylogeny of the chosen organisms (Figure 4). The position of AmphiS19 was intermediate between the vertebrate S19 and the invertebrate S19. This is in line with the notion that amphioxus is an organism transitional from invertebrate to vertebrate. Within the chordate clade, the AmphiS19 branched out first, forming an outgroup of the vertebrate S19. In addition, no particular S19 proteins in vertebrates showed consistently higher sequence similarity to AmphiS19 at the amino acid level. These data suggest that the AmphiS19 gene possibly represents the archetype of the vertebrate S19 genes.

Osmium isotopic compositions in Ganga river sediments

G. J. Chakrapani1, G. Gaillardet1, B. Dupré1 and C. J. Allegre2

1Laboratoire de Geochemie et Cosmochimie, Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris Cedex 05, France
2LMTG, University Paul Sabatier, 38 rue des 36 Ponts, Toulouse 31400, France

The Ganga river originates and flows through the Himalayas and then over the plains carrying enormous amount of sediment to the world’s oceans. The suspended sediments from selected locations in the entire river basin are analysed for osmium isotopic composition and reported in this study. The sediments are characterized by high Re/Os ratios and are extremely radiogenic as evident from their $^{187}$Os/$^{188}$Os isotopic ratios. Samples from the tributaries Alaknanda, Bhagirathi, Gandak and Ghaghra show pronounced $^{187}$Os/$^{188}$Os. The integrated effect is seen at Farakka, the farthest downstream location of the Ganga in the present study. High Os concentrations combined with sediment flux makes the Ganga an important source for soluble Os isotopic evolution in the oceans.

Interest in rhenium–osmium (Re–Os) systematics in rivers has risen sharply in recent years due to the revelation of changes associated with sea water Os isotopic compositions during the past 70 Ma. Radiogenic $^{187}$Os is produced from the $\beta$-decay of $^{187}$Re with a half-life of 4.2 x 10$^{10}$ years. Osmium isotopic composition in sea water is derived from the weathering of basaltic and peridotitic oceanic crust, hydrothermal solutions, additions from cosmic dust and continental weathering products. The $^{187}$Os/$^{188}$Os ratios of submarine weathering and cosmic inputs are nearly identical (~0.13) and about ten times lower than that of average continental matter (~1.26). Because of the large isotopic variations between these different sources, Os isotopes in the oceans convey the then prevalent continental weathering processes. The osmium isotopic composition of the present-day sea water is markedly higher since the past 70 Ma, similar to the $^{87}$Sr/$^{86}$Sr ratios in sea water. This enhancement in radiogenic Os in sea water is largely attributed to the Himalayan tectonics with its accompanying increased silicate weathering and in particular, chemical weathering of the extremely radiogenic black shales in the lesser Himalayan region.3,4 Because of its low concentration and high ionization potential (8.7 eV), Re–Os isotopic measurements are carried out by Negative Thermal Ionization Mass Spectrometry (NTIMS)$^5$.

It is estimated that annually about 1.6 x 10$^{16}$ g of sediment are transported to the oceans by rivers, approximately 10% of which is contributed by the Ganga–Brahmaputra river system. The drainage basin of the Ganga occupies an area of 1.06 x 10$^6$ km$^2$; while more than 60% of water flowing into the Ganga comes from the Himalayan source, 40% is contributed from the peninsular region. Briefly, lithology of the basin comprises7 Kumaun Himalayas consisting of Siwalik sediments composed of coarse sandstone, clays and conglomerates, the central–lower Himalayas of Krol formation consisting of dolomitic limestones, shales, quartzites, granites and gneisses, and the lower reaches of the basin characterized by alluvial plains which consist of massive beds of clay, sand and gravel with extensive calcareous concretions and saline/alkaline soils. The suspended sediments in the river are mostly medium-to-coarse silt (mean size <4.5 to 5.75 $\phi$) and are poorly sorted. The clay minerals abundant in the sediments are mica followed by kaolinite in the upstream and mica followed by smectite in the downstream region. Sedimentological and Sr isotopic data along with clay mineral composition of ODP Leg 116 and DSDP Leg 22 indicate the dominance of Ganga river sediments in the Bay of Bengal.$^8$

Suspended sediments from ten locations shown in Figure 1, covering the mainstream of the Ganga and all major tributaries were collected from 5 to 10 litres of water samples using sartorius 0.2 $\mu$m cellulose acetate filters and tangential filtration. Suspended sediments were then separated by centrifugation and desiccation at 80°C to remove the remaining water from the sediments. Re–Os isotopic measurement in sediments was carried out following the chemical procedures described by Birck et al.$^9$

ACKNOWLEDGEMENTS. We thank Drs Jianwen Wei and Chunyan Mou, College of Life Science Zhongshan University for help in cDNA library construction. We are also grateful to the anonymous reviewers for their valuable comments. The work was supported by a grant (2002 AA 629190) from the Ministry of Science and Technology, China to S.Z.

Received 17 June 2002; revised accepted 8 October 2002

*For correspondence. (e-mail: gjc3@iitr.ernet.in)