Structural and Phylogenetic Profiles of Muscle Actins from Cephalopods

Y. Ochiaia,c,*, S. Watabeb,c and G. Wangc

aDepartment of Marine Science and Technology, Tokai University, Shizuoka 424-8610, Japan
bKitasato University School of Marine Biosciences, Sagamihara, Kanagawa 252-0373, Japan
cFormerly, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan

Abstract: Structural properties of cephalopod muscle actins have not been characterized in detail to date. Sequences of these actins have been determined by cDNA cloning from the arm muscles of four cephalopod species, namely, Japanese common squid Todarodes pacificus, spear squid Loligo bleekeri, common cuttlefish Sepia esculenta, and common octopus Octopus vulgaris. The sequence identities of the deduced amino acid sequences among the cephalopod actins were around 98%, and the substitutions were mostly limited to Subdomains 1 and 3. Phylogenetic analyses revealed that cephalopod actins formed one cluster and were distinguishable from the bivalve counterparts. The modeled tertiary structure of squid actin was very similar to that of rabbit counterpart, but the structures of Subdomains 1 and 3 were slightly different.

Keywords: Actin, cephalopod, muscle, structure, sequence.

INTRODUCTION

Actin, a globular protein consisting of about 375 amino acids, is ubiquitous in organisms, and is involved in many cell motility and functions, namely, muscle contraction, sarcoplasmic streaming, cell division, transcription and cell shape regulation, etc [1, 2]. Actin is characterized by its so many target substances including actin binding proteins [2]. In striated muscles, actin is a major component, second to myosin, another contractile partner protein, and forms a thin filament together with the regulatory tropomyosin and troponin complex. Physiological role of actin in muscle is to activate myosin Mg²⁺-ATPase activity in the presence of Ca²⁺ ion [3]. Actin exists as a monomer (globular or G-actin) under low salt concentration, but polymerizes to form a double helical filament (filamentous or F-actin) under physiological ionic strength through a considerable conformational change [4].

G-actin is relatively a flat molecule of about 55×55×35 Å. The first high-resolution crystal structure of actin was provided by Kabach et al. [5] as a 1:1 complex with DNase I. Actin molecule is composed of two major (small and large) domains separated by a deep cleft, which binds a nucleotide and a divalent cation. More precisely, two clefts (upper and lower) formed between the two domains bind these ligands.

The steady-state mechanism of actin polymerization/depolymerization can be explained by treadmilling of actin filament. The bound nucleotide in a monomer is, in most cases, ATP. During polymerization, ATP is hydrolyzed to ADP [6]. The small and large domains are further divided to Subdomains 1 and 2, and Subdomains 3 and 4, respectively [7]. The lower cleft between Subdomains 1 and 3 is lined by predominantly hydrophobic residues such as isoleucine and leucine and mediates longitudinal contacts between actin subunits in F-actin. The cleft, designated as the target binding cleft, also forms the scaffold for the most actin binding proteins.

Thin filaments of vertebrates and invertebrates show only small differences in structure unlike thick filaments [8]. Actin is known for its highly conserved amino acid sequence, in spite of the fact that no less than 150 actin binding proteins of clearly different properties have been found [2]. The evolutionary rate of actin genes are considered to be very low [9]. However, even a few amino acid substitutions could cause a dramatic change in the function and stability of proteins as can be seen in deep-sea fish actin [10].

Regarding invertebrate actins, too little information is available to date. Among marine aquatic invertebrates, cephalopods such as squid and octopus possess well developed mantle muscles, and thus are fast swimmers in the sea. Cephalopods are the most complex class of mollusks, and are among the most highly evolved marine invertebrates [11]. It follows that they should have specialized locomotion system
including the contractile proteins, actin, myosin, etc. Such a background prompted us to characterize their actins which could be involved in their unique locomotion activities. In the present study, actin genes were cloned from four species of cephalopods, and their structures, from the primary to tertiary ones, were investigated in detail.

**MATERIALS AND METHODS**

**Materials**

Live specimens of Japanese common squid *Todarodes pacificus* (body weight: 303 g, full length 38 cm), spear squid *Loligo bleekeri* (317 g, 46 cm), and common cuttlefish *Sepia esculenta* (267 g, 21 cm) were slaughtered at Tsurugisaki, Kanagawa Prefecture. A live specimen of common octopus *Octopus vulgaris* (430 g, 38 cm) was slaughtered on site at Tokyo Metropolitan Central Fish Market just after purchasing. The arm muscles were dissected, and the skins were carefully removed. The muscles were cut into small pieces by sterilized scissors and immediately kept in RNAlater (QIAGEN) until used.

**cDNA Cloning**

Total RNA was extracted from the above samples. Total RNA extraction was performed by using ISOGEN (Nippon Gene) according to the manufacturer’s protocols with a little modification. The muscles were homogenized with 1 ml of Isogen in 1.5 ml tubes. The tubes were then swirled using vortex and incubated at room temperature for 5 min. 0.2 ml of chloroform was added into the tubes and the tubes were shaken vigorously for 15 sec and stored at room temperature for 2 to 3 min. After 2 min incubation at room temperature, the tubes were centrifuged at 16000 x g for 15 min at 4°C. The uppermost aqueous layers containing the extracted RNA were transferred into new tubes, added with 0.5 ml of isopropanol, and stored at -20°C overnight to precipitate RNA. The precipitated RNA was collected by centrifugation at 16000 x g for 30 min at 4°C. All the aqueous phase was discarded and the precipitated RNA was washed with 1 ml of 70% ethanol. The tubes were again centrifuged at 4600 x g for 10 min at 4°C. Alcohol was discarded and the tubes were dried briefly until the alcohol residue was completely evaporated. Finally, the RNA pellets were resuspended into 25 μl sterile distilled water. The concentration of each sample was determined spectrophotometrically.

Cloning of actin genes was carried out as shown in Figure 1 by combinations of six primers, four of which had been designed based on the conservative sequences of squid *L. pealei* actin. The first strand cDNA synthesis was performed by using SuperScript III Reverse Transcriptase (Invitrogen). The solution containing 5 μg RNA was diluted into 7 μl with distilled water, and mixed with 5 μl of 10 μM AP primer and 1 μl of 10 μM dNTP. The mixture was incubated at 65°C for 5 min and then at 0°C for 2 min. For the synthesis, the mixture was subsequently mixed with 4.0 μl of 5x First Strand Buffer, 1 μl of 0.1 mM dithiothreitol (DTT) and 1 μl of SuperScript III Reverse Transcriptase, and kept at 48°C for 1.5 h and then transferred into 70°C for 15 min. The solution was digested by 1 μl of RNase H (Invitrogen) at 37°C for 30-60 min. A PCR-based cloning strategy was used. Initially several pairs of degenerate primers were designed according to the conserved regions of the available actin cDNA sequences from the database. Other degenerate primers and specific primers were designed according to the subsequent sequencing results. For the 3’ end of the actin gene, a universal AUAP primer was used as the reverse primer.

By using these primers, PCR was performed using *Ex Taq* DNA polymerase (Takara). The reaction

![Figure 1: Strategies of cloning. CDS and UTR stand for the coding sequence and untranslated region, respectively. Three overlapping cDNA fragments for each species were cloned and sequenced. The dashed portions indicate those determined only for Japanese common squid and common octopus. GeneRacer RNA oligo, AUAP primers, and other specific primers designed based on the conservative sequences of spear squid *Loligo pealei* actin (AY701849) were used.](image-url)
mixture with the total volume of 20 μl contained 1 μl of first strand cDNA as a template, 2 μl of 10 x Ex Taq Buffer, 1.6 μl of 2.5 mM dNTP, 1 μl of 10 μM forward primer, 1 μl of 10 μM reverse primer, 0.2 μl of Ex Taq and 13.2 μl of distilled water. PCR consisted of 30 cycles of denaturation at 94°C for 5 s, annealing at 55-62°C according to the melting temperatures of each pair of primers for 30 s, and extension at 72°C for 1 min. The last cycle employed an extension of 5 min. For purification, the amplified PCR products were precipitated by the addition of 50 μl of 95% ethanol and 0.1 volume of 3 M sodium acetate (pH 4.8) and centrifugation at 15000 rpm for 30 min. The supernatant was then removed, and the pellets were washed with 50 μl of 70% ethanol and centrifuge at 15000 rpm for 10 min. Finally, the pellets were resuspended into 10 μl of distilled water. Three μl of the purified PCR products were subjected to 2% agarose gel electrophoresis, and the bands were visualized with ethidium bromide or SYBR Green (Invitrogen) staining.

Sequencing

The PCR products containing target bands were subcloned into pGEM-T vectors using pGEM-T Vector Systems (Promega) according to the manufacturer’s instructions. The reaction tubes were incubated overnight at 4°C to have the maximum number of transformed cells. The ligation products were then transformed into *Escherichia coli* JM 109 competent cells. The bacteria were subsequently plated onto agar plates and incubated at 37°C overnight. The individual recombinant white clones were picked up and subjected to insert-check PCR amplification. The PCR products were run on 2% agarose gel. The positive inserted clones were grown in LB liquid medium containing ampicillin overnight at 37°C with agitation.

Isolation and purification of inserted plasmid DNA from the cells were performed using GenElute Plasmid Miniprep Kit (Sigma) according to the manufacturer’s protocols. Prior to sequencing, the purified plasmid DNA was subjected to labeling PCR using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and sequencing was carried out using a genetic analyzer (ABI Prism 3100).

The sequences were aligned by CLUSTAL W. The accession numbers of amino acid sequence in the database referred in the present study were as follows: nematode *Caenorhabditis elegans* (P10983), freshwater snail *Biophalaria glabrata* (P92179), sea hare *Aplysia californica* (U01352, AY703989), mussel *Mytilus galovincialis* (AF157491), scallop *Placopecten magellancis* (Q26065), pearl oyster *Pinctada fucata* (EU726273), spear squid *Loligo pealei* (AY701849), brine shrimp *Artemia* sp. (P18600), Pacific transparent sea squirt *Ciona savignyi* (O15998), perch *Perca flavescens* (AY332493), zebrafish *Danio rerio* (AF180887), carp *Cyprinus carpio* (P53479), chicken *Gallus gallus* (AF012348, P68139), mouse *Mus musculus* (AA061020), rabbit *Oryctolagus cuniculus* (P68135-1), and human *Homo sapiens* (P68032).

Bioinformatic Analyses

The secondary structure prediction was performed using the GOR program, assuming that the predicted structure is basically one of the highest probabilities, compatible with a predicted α-helix segment of at least four residues. Hydropathy plot was carried out according to Kyte and Doolittle [12]. The tertiary structure of actin was obtained by using a program MolFeat ver.4.5 (FiatLux).

RESULTS

Primary Structures

By using the first strand cDNAs as the templates, three cDNA fragments were amplified by PCR using degenerated primers designed according to the conserved regions of the mollusk actin cDNA sequences obtained from GenBank database. For the 5’ end, several degenerated primers were designed from the start codon ATG, according to the conserved regions of the published mollusk actin cDNA sequences. By compounding these overlapping sequences, the full length actin cDNA sequences could be obtained.

The full length cDNA sequences of *T. pacificus*, *L. bleekeri*, *S. esculenta* and *O. vulgaris* actins consisted of 1280, 1406, 1469 and 1248 bp, respectively. All the cDNAs contained an open reading frame of 1131 nucleotides, encoding 376 amino acids (Figure 1). These sequences have been deposited to the database with the accession numbers as follows: FJ611946 for *T. pacificus*, GU338003 for *L. bleekeri*, GU338004 for *S. esculenta*, and FJ611947 for *O. vulgaris*.

The deduced amino acid sequences of actins from the above four species are aligned in Figure 2. The sequence identities of *T. pacificus* actin against *L. bleekeri*, *S. esculenta* and *O. octopus* actins were
98.4%, 97.9%, and 98.1%, respectively. The substitutions were limited to the middle region of Subdomain 1 as well as Subdomain 3, while those of Subdomains 2 and 4 were found to be essentially the same (only one conservative substitution between Todarodes and Octopus actin sequences). Including the reported sequence of Loligo pealei actin, the amino acid sequences of all the cephalopod muscle actins are highly conserved. Sepia actin was featured by the two substitutions, Gly109 and Leu122, both of which occurred in Subdomain 1, compared to the other three cephalopod actins.

In Figure 3, the sequences of actins from some vertebrates and invertebrates were aligned with the cephalopod (T. pacificus) actin. The identities of the sequences among actins from vertebrates and invertebrates were shown in Table 1. Based on the sequence identity, the cephalopod actins were closer to arthropod actin, but remote from vertebrate and sea squirt actins. Invertebrate actins were one residue shorter compared to the vertebrate actins, because there was one deletion at the N terminus (the fourth residue from the N terminus). Vertebrate actins were featured by the presence of Cys12, while invertebrate actins were featured by the presence of Gly109 and Leu122.

**Table 1**: Sequence identities of muscle actins from vertebrates and invertebrates.

<table>
<thead>
<tr>
<th>Actin Source</th>
<th>Sequence Identity</th>
<th>Cephalopod Actin</th>
<th>Vertebrate Actin</th>
<th>Invertebrate Actin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squid</td>
<td>98.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nematode</td>
<td>97.9%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shrimp</td>
<td>98.1%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabbit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carp</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Amino acid sequences of cephalopod muscle actins aligned on the basis of Japanese common squid Todarodes pacificus actin sequence. Loligo, Sepia and Octopus correspond to Loligo bleekeri, Sepia esculenta and Octopus vulgaris, respectively. Identical residues are indicated by dots. Boldfaced numbers represent the Subdomains 1 through 4, which are also indicated by the horizontal bars. The numbers on the right side indicate the residue numbers.
ones were characterized by Cys19. Though some substitutions were recognized in Subdomains 2 and 4, they were essentially conservative ones (Figure 3).

Based on the sequences, molecular weights and isoelectric points (pI) were calculated (Table 2), in comparison with the other actins. Molecular weights of invertebrate actins (including cephalopod ones) were smaller than the vertebrate (carp and rabbit) counterparts, except for sea squirt actin. Squid (T. pacificus) actin was found to have the same pI (5.22) with those of carp actin and was close to rabbit and sea squirt actins. On the other hand, octopus actin showed the same pI (5.30) with nematode and brine shrimp actins.

Secondary Structure

The structure predicted using GOR is shown in Figure 4. By referring to Figure 2, Subdomains 1 and 3 are rich in stranded sheets, and the Subdomain 4 is highly helical. The mid region of Subdomain 1 was also rich in the helical content. The ratio of α-helix, extended strand and random coil were estimated to be 29.5, 23.7 and 46.8%, respectively, and no other secondary structure was expected.

Phylogenetic Analysis

The phylogenetic tree constructed by a neighbor-joining method is shown in Figure 5. Cephalopod actins formed one cluster, and were clearly separated from vertebrate muscle actins. Cephalopod actins were closer to the other mollusk counterparts as expected. Bivalve actins showed unexpected variations from the molluskan ones.

Tertiary Structure

Hydropathy plot of squid (T. pacificus) actin is shown in Figure 6. Hydrophobic and hydrophilic regions appeared by turns. However, the latter took over the former, resulting in the overall hydrophilicity of the molecule. The most hydrophilic parts appeared in

---

Table 1: Identity of Amino Acid Sequences (%)

<table>
<thead>
<tr>
<th></th>
<th>Squid</th>
<th>Nematode</th>
<th>Shrimp</th>
<th>Squirt</th>
<th>Carp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squid</td>
<td></td>
<td>96.3</td>
<td>94.7</td>
<td>91.5</td>
<td>92.0</td>
</tr>
<tr>
<td>Nematode</td>
<td>96.3</td>
<td></td>
<td>96.0</td>
<td>91.5</td>
<td>96.0</td>
</tr>
<tr>
<td>Shrimp</td>
<td>94.7</td>
<td>96.0</td>
<td></td>
<td>91.8</td>
<td>96.0</td>
</tr>
<tr>
<td>Squirt</td>
<td>91.5</td>
<td>92.0</td>
<td>91.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carp</td>
<td>92.0</td>
<td>93.4</td>
<td>91.8</td>
<td>96.0</td>
<td></td>
</tr>
<tr>
<td>Rabbit</td>
<td>92.3</td>
<td>93.4</td>
<td>91.8</td>
<td>96.3</td>
<td>98.7</td>
</tr>
</tbody>
</table>

Table 2: Molecular Weight and Isoelectric Point of Actins

<table>
<thead>
<tr>
<th></th>
<th>Squid</th>
<th>Octopus</th>
<th>Nematode</th>
<th>Shrimp</th>
<th>Squirt</th>
<th>Carp</th>
<th>Rabbit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mol. Wt.</td>
<td>41708</td>
<td>41761</td>
<td>41796</td>
<td>41786</td>
<td>42103</td>
<td>41959</td>
<td>42051</td>
</tr>
<tr>
<td>pI</td>
<td>5.22</td>
<td>5.30</td>
<td>5.30</td>
<td>5.30</td>
<td>5.23</td>
<td>5.22</td>
<td>5.23</td>
</tr>
</tbody>
</table>

Figure 4: Secondary structure prediction of squid T. pacificus actin. The helical regions are indicated by the letter “h” together with the underlines. The letters “c” and “e” stand for the random coil and extended sheet, respectively. The structure was predicted by GOR.
Subdomains 1 and 2. The tertiary structure of *T. pacificus* actin was modeled based on the amino acid sequence (Figure 7). In the Figure, the structure of rabbit actin was superimposed. The structures of Subdomains 2 and 4 matched almost completely, while those of Subdomains 1 and 3 differed from each other only slightly.

**DISCUSSION**

All the sequenced actins in this study consisted of 376 amino acids (Figure 2). However, it has been reported that, after the translation, N-terminal Met and Cys residues are acetylated and cleaved, and then N-terminal Asp is reacetylated [13]. Partial amino acid substitutions in actins have also been reported for

---

**Figure 5:** Phylogenetic tree drawn by a neighbor-joining method based on the amino acid sequences of the vertebrate and invertebrate actins. Refer to the text for the accession numbers of the sequences.

**Figure 6:** Hydropathy plot of squid *T. pacificus* actin. The plot was performed according to Kyte and Doolittle [12]. The window size was set to 7.
several fish species [14]. The authors claimed the substitutions are responsible for thermal stability and resistance against urea as well as susceptibility to proteolysis. Mudalige et al. [15] also reported the effects of substitutions at the critical residue(s) in actin for its stability and myosin activation profiles. By referring to the substitutions observed among the actins examined (Figure 3), it was suggested that Subdomains 1 and 3 modulate the interaction with myosin.

Identity or high sequence similarity of Subdomains 2 and 4 are supported by the fact that they are the products of gene duplication [13]. Tropomyosin is located along the grooves formed by Subdomains 3 and 4, while part of Subdomain 1 is involved in myosin binding [7]. N and C termini as well as DNase I binding loop are located in Subdomain 1. McKane et al. reported the critical effects of mutations in Subdomains 1 and 2 of actin [16].

The phylogenetic tree obtained in the present study is reminiscent of the one reported by Hooper and Thuma [17], but shows more detailed relationship among invertebrate actins. In the present study, actin genes were cloned only from the muscles of cephalopods, and thus the multi form of genes as reported by Hooper and Thuma could not be detected in the species examined [17]. On the other hand, Carlini et al. [11] used the partial sequences of actin gene from 44 species of cephalopod as phylogenetic markers, and concluded that at least three actin loci exist, resulting in the expression of three isoforms, two of which are muscular type and another is cytoplasmic type.

Similar reports have been reported for sea hare Aplysia californica [18, 19], scallop Placopecten magellanicus [20], and abalone Haliotis iris [21]. In the case of lobster Homarus americanus, twelve genes encoding actin were detected, namely, eight from skeletal muscle, one from cardiac muscle and three from non-muscle cells [22]. The presence of isoforms in other invertebrates has been described by several groups [17, 23, 24]. On the other hand, Venkatesh et al. [25] characterized nine actin genes in puffer Takifugu rubripes. Mudalige et al. [15] also reported the presence of actin isoforms in the slow skeletal muscle.
of trout. Mercer et al. [14] demonstrated the presence of isoform solely in the smooth muscle of some fish species. Lancioni et al. [26] revealed the presence of actin isoforms in human satellite cells. Mammalian and avian actins consist of six isoforms encoded by different genes [1]. Although the sequence identities of these isoforms are reported to be higher than 93%, they are supposed to have different specialized functions.

CONCLUSION

Cephalopod actins were found to be very similar to those from other sources as considered from the high sequence identities (>91.5%), but were clearly distinguishable even from the bivalve actins in the phylogenetic tree. The amino acid substitutions (up to around 40 residues) might be responsible for the effective contractility of cephalopod muscles, facilitating quick locomotion in water. Even one amino acid substitution would greatly affect the biological functions of actin. Point mutation and kinetic studies using reconstituted actomyosin would be useful for demonstrating this hypothesis.

ACKNOWLEDGEMENTS

The expenses of the present study were defrayed in part by a Gran-in-Aid from Ministry of Education, Science, Sport and Technology of Japan (KAKENHI #19380119 to YO).

REFERENCES

