Two Divergent Isotypes of the Fourth Complement Component from a Bony Fish, the Common Carp (Cyprinus carpio)

Junichi Mutsuro, Noriyuki Tanaka, Yoko Kato, Alister W. Dodds, Tomoki Yano and Miki Nakao

*J Immunol* 2005; 175:4508-4517; doi: 10.4049/jimmunol.175.7.4508
http://www.jimmunol.org/content/175/7/4508

**References**

This article cites 46 articles, 12 of which you can access for free at:
http://www.jimmunol.org/content/175/7/4508.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Two Divergent Isotypes of the Fourth Complement Component from a Bony Fish, the Common Carp (Cyprinus carpio)¹,²

Junichi Mutsuro,* Noriuki Tanaka,* Yoko Kato,* Alister W. Dodds,† Tomoki Yano,* and Miki Nakao³*

Duplication and diversification of several complement components is a striking feature of bony fish complement systems. It gives an interesting insight into an evolutionary strategy for the possible enhancement of the repertoire of innate immunity. The present study is aimed at examining diversity in bony fish C4, a member of the thioester-containing complement components. Two diverged cDNA sequences sharing only ~32% identity at the amino acid level were isolated from the common carp and designated C4-1 and C4-2. C4-1 and C4-2 share a number of C4-like structural signatures, such as the thioester site and a disulfide-linked three-chain structure. Interestingly, they differ at the residue corresponding to the thioester-catalytic histidine, as seen in the human C4A and C4B isotypes, suggesting their distinct substrate specificities in the binding reaction of the thioester. Phylogenetic analysis indicates that the divergence of C4-1 and C4-2 predated the separation of the cartilaginous and bony fish lineages. Genomic Southern hybridization suggests the presence of single copy genes encoding C4-1 and C4-2 in the carp genome. An activation fragment, C4a, was shown to be released from each isotype in carp serum activated via the classical and/or lectin pathways. Synthetic peptides representing a putative C2 binding site on C4-1 and C4-2 inhibited the classical pathway-mediated hemolytic activity of carp serum in a dose-dependent manner. The results suggest that C4-1 and C4-2 represent two major lineages of C4 that are present in carp serum, have distinct binding specificities, and are functional in the classical/lectin pathways of complement activation. The Journal of Immunology, 2005, 175: 4508–4517.
the complement components. In the present study, we have isolated two diverged cDNA clones representing distinct C4 isotypes from a tetraploid bony fish, the common carp (Cyprinus carpio). Unlike C3, C5, and \( \alpha \)-M that show 75–90% identity among the isotypes, the two carp C4 sequences share only 32% amino acid identity. Interestingly, one of them lacked the catalytic His as does human C4A, suggesting their difference in the binding specificity through the thioester. The functional significance of the two carp C4 molecules is also discussed.

Materials and Methods

Materials

Restriction enzymes and ISOGEN were purchased from Nippon Gene. Moloney murine leukemia virus (MMLV)-reverse transcriptase (RTase), oligo(dT) primer, PCR SuperMix High Fidelity and TA cloning kit were obtained from Invitrogen Life Technologies, and pGEM-T vector was obtained from Promega. PCR DIG Probe Synthesis kit, DIG blocking reagent, and alkaline phosphatase-conjugated anti-digoxigenin (DIG) Ab (Fab) were purchased from Roche Molecular Biochemicals. Synthetic oligonucleotides were obtained from Hokkaido System Science. The common carp weight (2M) that show 75–90% identity among the carp hepatopancreas was screened by plaque hybridization with DIG-labeled probes as described elsewhere (27). Positive clones were plaque-purified and subcloned into pBluescript SK(−) by in vivo excision according to the manufacturer’s instructions (Stratagene).

RT-PCR amplification of carp C4 cDNA segments

Total RNA was purified from carp hepatopancreas as described elsewhere using ISOGEN reagent (25). Oligo(dT)-primed first-strand cDNA was synthesized using MMLV-RTase from the total RNA and subjected to PCR amplification of a carp C4-like cDNA sequence using a set of primers, P1 and P2 (Table I), which correspond to well-conserved amino acid sequence stretches of mammalian C4. The PCR-amplification was conducted using a TP-3000 thermocycler (Takara), under the following conditions: 30 cycles of 95°C for 0.5 min, 50°C for 0.5 min, and 72°C for 1 min.

For amplification of another carp C4-like sequence, the total RNA was reverse-transcribed using MMLV-RTase and an antisense primer P3 (Table I), which corresponds to a conserved amino acid sequence, IDIQAEM, in the N-terminal region of carp C4 α-chain isolated from serum, and P5 corresponds to a conserved sequence, FPQSWMW, found at its 60 residues C-terminal in the carp C4-1 sequence. Amplified products were gel-purified and subcloned into pGEM-T vector for sequencing.

Colony hybridization

Colonies grown on Luria-Bertani/amp plates were transferred in duplicate to Hybond N membranes, followed by cell lysis, denaturation of the DNA, and neutralization, as described elsewhere (26). Hybridization was performed with DIG-labeled cDNA probes under stringent conditions, according to the manufacturer’s instructions (Roche Diagnostic Systems). Positive colonies were visualized with alkaline phosphatase-conjugated anti-DIG using 5-bromo-4-chloro-3-indolylphosphate (X-phosphate) and NBT in vol of water and applied to a Q-Sepharose FF column (1.6 cm) equilibrated with 2 M ammonium acetate, 0.5 M NaCl. Positive colonies were visualized with alkaline phosphatase-conjugated anti-DIG using 5-bromo-4-chloro-3-indolylphosphate (X-phosphate) and NBT as substrates.

Library screening

Once amplified, a cDNA library constructed using λZAP II vector from carp hepatopancreas was screened by plaque hybridization with DIG-labeled probes as described elsewhere (27). Positive clones were plaque-purified and subcloned into pBluescript SK(−) by in vivo excision according to the manufacturer’s instructions (Stratagene).

Nucleotide sequence analysis

The nucleotide sequences were determined from both strands by the dideoxy chain-termination method (28) using a model 377 sequencer and Big-Dye Terminator Cycle Sequencing kit (Applied Biosystems).

Southern hybridization

Genomic DNA (5 μg) isolated from carp erythrocytes as describe elsewhere (26) was digested to completion at 37°C for 16 h with 50 U of BanHI, EcoRI, HindIII, or SacI. The digests were electrophoresed on a 1% agarose gel, transferred to Hybond N* membranes (Amersham Biosciences), and UV cross-linked. Hybridization and washing under stringent conditions and chemiluminescent detection of DIG-labeled probes were performed as described elsewhere (27).

Phylogenetic analyses

Amino acid sequences of the thioester-containing proteins were obtained by BLASTP search of the SwissProt, TrEMBL, and TrEMBL_new databases using the carp C4-2 sequence as a query and aligned using the ClustalW 1.8 software (29) under the following parameters: gap opening, 10.0; gap extension, 0.2; delay divergent sequences, 30%; DNA transition weight, 0.5; and protein weight matrix, Gonnet 250. On the basis of the alignment, phylogenetic trees were constructed using the neighbor-joining method (30) and the maximum likelihood method (31), with the aid of the MOLPHY software package (32). In both methods, all the gaps generated in the alignment were omitted for the analysis, and multiple substitution was corrected. Statistic significance of each branch was examined by the boot-strap percentages obtained from 1000 replications.

Purification of carp C4 proteins

Carp C4 was purified from serum by a modification of the method described for human C4 (33). Briefly, serum was made 5% with polyethylene-glycol 4000. After centrifugation, the supernatant was diluted with 2 vol of water and applied to a Q-Sepharose FF column (1.6 × 5 cm) equilibrated with 20 mM Tris-HCl buffer (pH 7.4) containing 100 mM NaCl, 50

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Direction</th>
<th>Encoding Amino Acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>GNTGDAQNGACARACAGATG</td>
<td>Sense</td>
<td>GCAEQTML(C4-1)</td>
</tr>
<tr>
<td>P2</td>
<td>ACRAANGCGNTAGCCA</td>
<td>Antisense</td>
<td>WLTAFTV(C4-1)</td>
</tr>
<tr>
<td>P3</td>
<td>AGCCCAAATCTCCTTTTGGAG</td>
<td>Antisense</td>
<td>TPSSIWL(C4-1)</td>
</tr>
<tr>
<td>P4</td>
<td>AATHGAYTHACRGCGATATG</td>
<td>Sense</td>
<td>IDIQAEM(C4-2)</td>
</tr>
<tr>
<td>P5</td>
<td>CCACATCCANGAYTNGRAA</td>
<td>Antisense</td>
<td>FPQSWMW(C4-1)</td>
</tr>
<tr>
<td>P6</td>
<td>GACCTTGTCGCCAGTGGAG</td>
<td>Antisense</td>
<td>SMRKPSC(C4-2)</td>
</tr>
<tr>
<td>P7</td>
<td>GAACCTTCTCTTTGGGOS</td>
<td>Antisense</td>
<td>PKQKVS(C4-2)</td>
</tr>
<tr>
<td>P8</td>
<td>CCCAACTCTGATCTCCCAACCC</td>
<td>Sense</td>
<td>AQSMQTC(C4-1)</td>
</tr>
<tr>
<td>P9</td>
<td>GGTGTACCTGTTGGGACATT</td>
<td>Antisense</td>
<td>NAQRAHC(C4-1)</td>
</tr>
<tr>
<td>P10</td>
<td>GCCAGAAGCATAGATGACG</td>
<td>Sense</td>
<td>EQAYER(C4-2)</td>
</tr>
<tr>
<td>P11</td>
<td>CGTCGATGAGCCAGACA</td>
<td>Antisense</td>
<td>VCLTDRC(C4-2)</td>
</tr>
<tr>
<td>P12</td>
<td>TACAGAAGAGGAGGCTTATC</td>
<td>Sense</td>
<td>QNERAYQC(S11)*</td>
</tr>
<tr>
<td>P13</td>
<td>TTTGCTACCTTCTAGAGACTT</td>
<td>Antisense</td>
<td>NVLKVTG(S11)*</td>
</tr>
</tbody>
</table>

a 40S ribosomal protein S11 subunit.
nM 6-aminocaproic acid, 5 mM EDTA, 0.02% NaN3, and 0.2 mM PMSF. After washing the column with 20 ml of the same buffer, adsorbed proteins were eluted with an NaCl gradient up to 500 mM (total 50 ml). The fractions (2 ml) were analyzed by SDS-PAGE under reducing conditions. C4-containing fractions were pooled, concentrated with Centricon CF-25, and passed through a TSKgel G3000 SWXL column (0.8 × 30 cm) equilibrated with the same buffer as that used for Q-Sepharose FF. Fractions rich in C4 were pooled and applied to a Mono Q 5/5 column equilibrated with the same buffer but containing 200 mM NaCl. The column was developed with a linear gradient up to 700 mM NaCl (total 30 ml). C4-containing fractions were pooled.

Protein analysis
SDS-PAGE and electroblotting onto polyvinylidene difluoride membranes were performed as described (34, 35). N-terminal amino acid sequence was determined on a model PPSQ-21 sequencer (Shimadzu Scientific Instruments).

Identification of C4a fragments
Carp serum (40 ml) was incubated with 10 mg/ml zymosan at 25°C for 1 h. After centrifugation, the supernatant was fractionated by gel filtration through a Cellulofine GCL-300-m column (4 × 60 cm) equilibrated with 10 mM sodium phosphate buffer (pH 7.0) containing 150 mM NaCl and 5 mM EDTA. Fractions corresponding to molecular mass of ~10 kDa were pooled, diluted 3-fold with 10 mM MES-NaOH buffer (pH 6.0), and loaded to a cation-exchange column of TSKgel SP-5PW (Tosoh) equilibrated with the same buffer but containing 75 mM NaCl, followed by elution with an NaCl gradient up to 600 mM. Fractions were analyzed by SDS-PAGE using Tris-tricine buffer system (36). The gel were reverse-stained with imidazole and zinc chloride (37), destained with 3% citric acid, and followed by electroblotting onto a polyvinylidene difluoride membrane for N-terminal sequencing.

Synthetic peptides
Two peptides, DDTLDEIVINEDAI and NSNEVEDFSDFGIET, representing carp C4-I and C4-2, respectively, and a peptide, DDYYTESEEIVSRTQ, corresponding to an equivalent position in carp C3-H1 and C3-S isomers were purchased from Sawady Technology. An acidic peptide, EEG using VNSINQATDDDS, served as a control. All the peptides have amidated C-termini, and they were purified by reverse-phase HPLC to achieve >90% purity by the supplier. Peptide stock solutions prepared in 5 mM imidazole and zinc chloride (37), destained with 3% citric acid, and followed by electroblotting onto a polyvinylidene difluoride membrane for N-terminal sequencing.

Hemolytic assay
Hemolytic activity of carp serum via the classical and alternative pathways were assayed essentially as previously described (38), using 96-well microtiter plates. Isotonic buffers used for the hemolytic assay were 2.5 mM barbital, 72.5 mM NaCl, 2.5% glucose, 0.1% gelatin, 0.15 mM CaCl2, and 0.5 mM MgCl2 (pH 7.5) (GGVB2EGTA, and 5 mM barbital, 145 mM NaCl, 10 mM EGTA, and 10 mM MgCl2 (pH 7.5) (Mg-EGTA-GBV). Sheep erythrocytes sensitized with carp Abs (EA) were prepared as described elsewhere (39). For the assay of the classical pathway-mediated hemolytic activity, 4 × 105 cells of EA were incubated with 60 µM of 1/60-diluted carp serum in a total volume of 150 µl at 25°C for 60 min, in the presence of 0–100 µM synthetic peptide. GGVB2EGTA was used as the diluent. After centrifugation, A540 nm of the supernatant was measured to evaluate the degree of hemolysis. The diluted serum gave ~30% hemolysis of EA input in the absence of the peptide.

The alternative pathway-mediated activity was assayed in a 150 µl reaction mixture composed of 4 × 105 cells of rabbit erythrocytes (Er), 60 µM of 1/20-diluted carp serum, and 0–100 µM peptide, using Mg-EGTA-GBV as a diluent. After incubation at 20°C for 2 h, the degree of hemolysis was measured as above. The diluted carp serum yielded 30–40% hemolysis of Er input in the absence of the peptide.

Relative hemolytic activity (RA) was calculated from the following formula: RA = Y/Y0, where Y and Y0 are fractions of cells lysed of test sample containing the peptide and of control sample lacking the peptide, respectively. Inhibitory activity of the peptides against the hemolytic activity via either pathway was evaluated as IC50, a peptide concentration that gives 50% decrease of the RA. For both target cells (EA and Er), negative control reactions in the presence of 10 mM EDTA were included.

Results
Purification of C4 protein from carp serum
Carp C4 was purified from serum by the four-step procedure composed of precipitation with 5% polyethylene glycol, Q-Sepharose FF chromatography, gel filtration on TSKgel G3000SWXL, and MonoQ chromatography (Fig. 1A). SDS-PAGE analysis of the final purification product revealed a human C4-like three polypeptide chain structure (α-chain, 95 kDa; β-chain, 66 kDa; γ-chain, 35 kDa) (Fig. 1B). N-terminal amino acid sequences of the three chains show close similarity with the corresponding region of medaka C4 and catfish C4 (data not shown). Some C4 preparations showed doublet bands of α-chain with the molecular masses of 95 and 93 kDa, both of which had an identical N-terminal sequence. This phenomenon has also been observed in a wild-type human C4 expressed and secreted from COS-1 cells (40).

Isolation of cDNA clones encoding carp C4
RT-PCR amplification of the thoeste-containing sequence of carp C4 using primers corresponding to the amino acid sequences GCAEQTM and WLTAFV yielded a single band with the expected size (~220 bp). The DNA was gel-purified and subcloned into pGEM-T vector for sequencing. As 36 randomly selected clones gave only C3 and α1M sequences, an additional 200 clones were screened by colony hybridization with a C3 probe, representing 2997–3646 of carp C3-H1-1 cDNA (21), and with an α1M

FIGURE 1. Purification of carp C4 from serum. Carp C4 was purified from serum by the four-step procedure composed of precipitation with 5% polyethylene glycol precipitation, Q-Sepharose FF chromatography, gel filtration on TSKgel G3000SWXL, and MonoQ chromatography (A). Elution profile of MonoQ chromatography (the final step). The post-gel filtration sample was applied to a MonoQ column and followed by gradient elution with NaCl. Carp C4 eluting at a major peak at ~500 mM NaCl was pooled as indicated by the bar. B, SDS-PAGE analysis of the purified carp C4. The purified C4 protein was run on a 10% gel under reducing conditions, and the gel was stained with Coomassie Blue R250. Molecular masses determined for α, β-, and γ-chains are shown on the left, and their N-terminal amino acid sequences are on the right.
probe, representing 2925–3136 of carp α2M-1 (23). Of the 20 clones that were negative to both probes, one clone (Cp1) showed closer similarity to human and Xenopus C4 than to C3 and α2M (Fig. 2). A carp hepatopancreas library containing $5 \times 10^5$ plaques was screened using Cp-1 as a probe. From 16 positive clones, 9 clones were sequenced from both termini, and a clone containing the longest insert (5.5 kbp), designated C4-1, was sequenced over its entire length. The nucleotide sequence of carp C4-1 begins with 24 bp of 5'-untranslated region and contains a 5100-bp long open reading frame encoding 1700 residues of proC4, followed by a 671-bp long 3'-untranslated region, including 22 bp of the poly(A) tail (DDBJ accession no. AB037278). Consistent with the three-chain structure observed for carp C4 protein, the deduced amino acid sequence contains two posttranslational processing signal sequences, RQKR and RRKR, at the same positions as those of human C4 (Fig. 3). However, the deduced N-terminal sequences of the chains were substantially different from those determined from the C4 protein, indicating the presence of an additional C4 isoform in the carp.

We next attempted to amplify a cDNA fragment that directly encodes the isolated carp C4 protein. A sense strand primer is based on the protein sequence of the α-chain, and an antisense strand primer corresponds to the amino acid sequence FPQSWMW, which is present in carp C4-1 and is substantially conserved among C4 sequences available in the databases. An amplified product with the expected size (~300 bp) was gel-purified and subcloned as above. Sequencing of 16 randomly selected clones yielded a clone, designated Cp-2, that was similar to but distinct from the corresponding region of carp C4-1 (Fig. 2). Rescreening of the library with Cp-2 probe detected 63 positive clones. Among these, 12 clones were randomly selected for sequencing from both termini. Ten had the same insert of 5.0 kbp in length, one of which, named C4-2, was sequenced for its entire length. C4-2 encodes an amino acid sequence containing the N-terminal sequences of the α- and γ-chains of the purified carp C4 protein but lacks the 5'-end sequence, including the region that should encode N-terminal 30 amino acids of proC4. The missing 5'-region of C4-2 was obtained by 5'-RACE. The complete C4-2 sequence contained 31 bp of 5'-untranslated region, an open reading frame of 5178 bp encoding 1716 aa, and 1164 bp of 3'-untranslated region, including a stop codon and a 22-bp long poly(A) tail (DDBJ accession no. AB037279). As in carp C4-1, two posttranslational processing signal sequences, RNKR and RRKR, were found in the C4-2 sequence (Fig. 3). The amino acid sequence identity between carp C4 sequences and other homologous proteins were calculated by pairwise alignment and shown in Table II.

**FIGURE 2.** RT-PCR amplification of cDNA segments encoding carp C4. Cp-1 and Cp-2 encoding C4-like sequences were amplified from the carp hepatopancreas RNA using degenerated primers corresponding to conserved amino acid stretches and N-terminal sequences of α-chain of carp C4 purified from serum. The deduced amino acid sequences of Cp-1 and Cp-2, along with mouse and medaka C4 sequences, are shown with a schematic model of proC4, which has a β-α-γ-chain organization from the N terminus.

**Multiple alignment of C4 sequences**

The deduced amino acid sequences of carp C4-1 and C4-2 were aligned with those of trout C4, fugu C4, medaka C4, shark C4, and human C4A, using ClustalW, version 1.83 software. As shown in Fig. 3, all the sequences possess the structural signature for C4: two posttranslational processing signal sequences that generate α-, β-, and γ-chains, the thioester site (CGEQ, CAEQ, or CALQ) in the middle of the α-chain, and a C1s/MASP2-cleavage site to generate C4a and C4b fragments. However, carp C4-1 and C4-2 show substantial sequence difference, sharing only 32% amino acid identity. Carp C4-2 is more similar to C4 of trout, fugu, and medaka than to carp C4-1 (Table II). An intriguing amino acid substitution is noted at the position corresponding to the catalytic His, which accelerates the binding reaction of the thioester to hydroxy-groups. Although carp C4-2 maintains the catalytic histidine in agreement with medaka and trout C4, the His is replaced by Asp in carp C4-1 as in human C4A and by Asn in Fugu C4, suggesting that the thioester of carp C4-1 and fugu C4 preferentially binds to amino-groups as does human C4A.

**Phylogenetic tree**

The molecular evolution of the α,M-family proteins (α2M, C3, C4, and C5) was analyzed using phylogenetic trees drawn by the neighbor-joining method and the maximum likelihood method. As shown in Fig. 4A (maximum likelihood method), carp C4-2 forms a cluster with medaka, fugu, and trout C4 in the cluster of vertebrate C4, whereas carp C4-1 and shark C4 form another cluster, which branches from the common ancestor of other C4 used here. This branching pattern is supported by high bootstrap percentages (97–100%). In the neighbor-joining tree, the cluster of carp C4-1 and shark C4 forms a larger cluster with the clade of the teleost C4 containing carp C4-2, although this clustering is supported by a low bootstrap percentage (63%). Taken together, the results suggest that the divergence of carp C4-1 and C4-2 predates the separation of carp from the other species in the teleost lineage, or it may be a much earlier event, and that carp C4-1 and shark C4 may represent a unique lineage of C4 with an ancient origin.

**Number of gene copies of carp C4-1 and C4-2**

To estimate the copy number of C4-1 and C4-2 genes in the carp genome, Southern hybridization was performed using cDNA probes representing the α- and β-chains of carp C4-1 and C4-2. The α-chain probes span putative exons 21–25 based on the exon-intron organization of fugu C4 gene (11). The β-chain probes of C4-1 and C4-2 correspond to the putative exons 3–4 and 3–8, respectively. No restriction site for the enzymes used here was
FIGURE 3. Alignment of the amino acid sequences of C4-1 and C4-2 from carp (Cyca) with those of trout (Onmy) C4, fugu (Taru) C4, medaka (Orla) C4, shark (Trsc) C4, and human (Hosa) C4A generated using the ClustalW software. Dots show residues that are the same as those of the carp C4-2 sequence, and dashes are gaps introduced for maximum matching. The residue number is given on the left of each sequence lane. The posttranslational processing sites between the /H9252/ - and /H9251/ -chains and between the /H9251/ - and /H9253/ -chain are denoted by overhead doublet lines. The thioester site, the catalytic His, and the Glu at the position /H11001/ are marked by “/H11001/”, “#”, and “$”, respectively. The amino acid sequences determined for the respective chains of carp C4 protein are in bold, with the N-terminal residue marked by slash. A backslash indicates the C1s/MASP-cleavage site. The overhead wave line corresponds to the C4d-specific loop region.
contained in all the probe sequences. As shown in Fig. 5A, the α-chain probes gave distinct band patterns for each digest, indicating that C4-1 and C4-2 are encoded by distinct genes. When HindIII digests from six representative carp were examined (Fig. 5B), the β-chain probes for C4-1 and for C4-2 each detected a single band, except for the number 6 fish, which gave two bands of C4-2, perhaps representing allelic variation in the C4-2 gene. These results suggest that C4-1 and C4-2 are encoded by single each genes in the carp genome.

Expression analysis by RT-PCR

The tissue distribution of mRNAs encoding C4-1 and C4-2 was analyzed by RT-PCR, using isotype-specific primers (P8/P9 for C4-1 and P10/P11 for C4-2; Table I) and total RNA from hepatopancreas, head kidney, body kidney, and spleen as templates. A cDNA fragment coding for the 40S ribosomal protein S11 subunit (a positive control) was also amplified using P12 and P13 (Table I) as described elsewhere (41). As shown in Fig. 6, both C4-1 and C4-2 are mainly expressed in the hepato pancreas.

Presence of two diverged C4 lineages in other teleost species

The presence of two diverged C4 lineages represented by carp C4-1 and C4-2 in other teleosts was examined by database-mining using the dbEST and Fugu Genome databases. TBLASTN searches against the dbEST database, using carp C4-1 and C4-2 sequences as queries, yielded both C4-1-like and C4-2-like cDNA sequences of zebrafish, rainbow trout, and Atlantic salmon as shown in Table III, implying that both the C4-1-like and C4-2-like lineages are also present at least in cyprinid and salmonid families in the teleost lineage. The expressed sequence tag (EST) sequence BI673185 in Table III homologous to carp C4-1 corresponds to a zebrafish genomic contig AL935181, which includes one of the two sequence segments encoding the zebrafish C4 sequence in a literature (9). In trout, the carp C4-2-like EST, CA362456, is 99% identical to the reported C4 mRNA sequence (10). In channel catfish, there is only one hit (BM438332) showing higher similarity to C4-2 than to C4-1 of carp. In contrast, TBLASTN searches of the Fugu genome database using carp C4-1 and C4-2 queries gave hits to an identical scaffold, M000795, that shows nearly equal similarity to carp C4-1 and C4-2. However, analysis of the complete sequence (Table II; Fig. 4) indicates that this

<table>
<thead>
<tr>
<th>Table II. Amino acid sequence identity (%) of carp C4-1 and C4-2 with C4 from other vertebrates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carp</strong></td>
</tr>
<tr>
<td>Carp C4-2</td>
</tr>
<tr>
<td>Human C4A</td>
</tr>
<tr>
<td>Human C4B</td>
</tr>
<tr>
<td>Mouse C4</td>
</tr>
<tr>
<td>Xenopus C4</td>
</tr>
<tr>
<td>Medaka C4</td>
</tr>
<tr>
<td>Trout C4</td>
</tr>
<tr>
<td>Fugu C4</td>
</tr>
<tr>
<td>Shark C4</td>
</tr>
</tbody>
</table>

**FIGURE 4.** Phylogenetic trees of the thioester-containing protein family drawn by the maximum likelihood method (A) and the neighbor-joining method (B). Bootstrap percentages after 1000 replications are given to relevant partitionings. Abbreviations: A2M, α2-macroglobulin; A1M, α1-macroglobulin; CVF, cobra venom factor; Slp, sex-limited protein; TEP, thioester protein; Anmi, spotted wolffish (Anarhichas minor); BrBe, amphioxus (Branchiostoma belcheri); Capo, guinea pig (Cavia porcellus); Ciin, ascidian (Ciona intestinalis); Ctid, grass carp (Ctenopharyngodon idella); Cyca, carp (Cyprinus carpio); Drme, fruit fly (Drosophila melanogaster); Epha, hagfish (Eptatretus burgeri); Gaga, chicken (Gallus gallus); Hosa, humans (Homo sapiens); Haro, Japanese solitary ascidian (Halo- cynthia roretzi); Laja, lamprey (Lampetra japonica); Mumu, mouse (Mus musculus); Nana, cobra (Naja naja); Onmy, trout (Oncorhynchus mykiss); Orla, medaka (Oryzias latipes); Paol, Japanese flounder (Paralichthys olivaceus); Rano, rat (Rattus norvegicus); Stpu, sea urchin (Strongylocentrotus purpuratus); Swex, coral (Swiftia exserta); Tara, fugu (Takifugu rubripes); Trsc, banded hound shark (Triakis scyllium); and Xela, Xenopus laevis.
followed by blotting to a Hybond N/Hi1001 HindIII (H), or C4-2 A probes corresponding to /Hi9252 and C4-2 (nucleotide no. 340–915).

Identification of C4a fragments generated from carp C4-1 and C4-2

To examine whether both carp C4-1 and C4-2 are involved in the classical and/or lectin activation cascades but not in the alternative pathway.

Identification of C4a fragments generated from carp C4-1 and C4-2

To examine whether both carp C4-1 and C4-2 are involved in classical/lectin pathway activation, synthetic peptides representing carp C4A and C4B were synthesized, containing protein family, such as C3, C5, and C4B, were tested for their ability to inhibit complement-mediated hemolytic reaction by carp serum. The C4-1 peptide, DDTLDEIVINEDAI, and C4-2 peptide, NSNEVEDFSDFGIET, correspond to the N-terminal region of the α’-chain of carp C4-1 (nucleotide no. 2629–3209) and C4-2 (nucleotide no. 340–915).

Inhibition of the classical pathway-mediated hemolysis by synthetic peptides representing carp C4A and C4B

To examine whether both carp C4-1 and C4-2 are involved in classical/lectin pathway activation, synthetic peptides representing the proposed C2 binding sites on C4-1 and C4-2 were tested for their ability to inhibit complement-mediated hemolytic reaction by carp serum. The C4-1 peptide, DDTLDEIVINEDAI, and C4-2 peptide, NSNEVEDFSDFGIET, correspond to the N-terminal region of the α’-chain that has been suggested as the C2 binding site of human C4 (40).

As shown in Fig. 8, both C4-1 and C4-2 peptides inhibited the classical pathway-mediated hemolysis in a similar dose-dependent manner, having an IC50 of ~60 μM. In contrast, neither peptide showed inhibition of alternative pathway-mediated hemolysis. The C3 peptide, on the other hand, inhibited the hemolytic reactions through both pathways with an approximate IC50 of 100 μM. A control peptide containing five acidic residues showed no significant interference. These results indicate that both C4-1 and C4-2 are involved in the classical and/or lectin activation cascades but not in the alternative pathway.

Discussion

In the present study, we have gained molecular evidence for the presence of two diverged isotypes of the complement component C4 in the common carp. Based on the number of chromosomes (e.g., the common carp, 2n = 100 (50 pairs of chromosomes); the tench Tinca tinca, 2n = 48; the barbel steed Hemibarbus labeo, 2n = 50; the grass carp Ctenopharyngodon idellus, 2n = 48 (data from FishBase, www.fishbase.org/)), the common carp is believed to have arisen by allotetraploidization 12–50 million years ago in the cyprinid family (42–44). Hence, it seems reasonable to assume that the tetraploidization has probably contributed, at least in part, to the gene duplications of other members of the thioester-containing protein family, such as C3, C5, and αM, of this species (21–23). However, it is unlikely that carp C4-1 and C4-2 were generated by the tetraploidization event as judged from their substantial sequence divergence. The phylogenetic tree (Fig. 4) also suggests that the C4-1 and C4-2 lineages arose earlier than the divergence of carp, trout, medaka, and fugu or possibly even earlier than that of bony and cartilaginous fish. Supporting this suggestion, database-mining using the dbEST and Fugu Genome databases indicated the presence of both C4-1-like and C4-2-like genes in zebrafish, rainbow trout, and Atlantic salmon (Table III). Furthermore, a C4 has been cloned recently from a shark, and it forms a cluster with carp C4-1 in the phylogenetic tree and is also carp C4-1-like in that it also lacks the catalytic histidine (12). Because a human C4B-like C4 with the catalytic histidine, rather
than C4A-like one lacking the histidine, has been considered to be a C4 form essential to host defense (45), the shark may have another C4 gene encoding a human C4B-like protein with the catalytic histidine, although it is yet to be sought. In contrast, a Fugu Genome database search with carp C4-1 and C4-2 queries implied that only a single C4 gene is present in the Fugu genome. Because the fugu (the tiger puffer, Takifugu rubripes) is a highly developed and specialized species in the phylogeny of teleosts, one of the two C4 lineages might have been lost somewhere in the teleost lineage after the separation of the common ancestor of fugu from that of carp and trout. Likewise, one of the two C4 lineages could have been lost in the vertebrate lineage leading to tetrapods before the divergence of amphibians, which apparently retain only one C4 gene, although at present it is not possible to say whether this is more closely related to the teleost C4-1 or C4-2 gene.

The most intriguing difference between carp C4-1 and C4-2 is that the catalytic histidine (His1123) present in C4-2 is replaced by an aspartate (Asp1109) in C4-1. As proven by ample data from a panel of artificial mutants of human C3 and C4 (46, 47), in which the catalytic histidine is substituted with less nucleophilic residues, carp C4-1 and C4-2 probably differ in the substrate specificity of the binding reaction of the thioester, i.e., C4-1 favors amino groups and C4-2 prefers hydroxyl groups. In addition to the catalytic histidine, a glutamate at its two residues C-terminal has been predicted to sharpen the substrate specificity. In human C3, the glutamate residue has been inferred to render the catalytic histidine with stronger nucleophilicity, providing a strict substrate specificity of the thioester toward hydroxyl-groups (48). Having a serine at the corresponding position, human C4B is reactive to both hydroxyl- and amino-groups (3, 49). Carp C4-2 and medaka C4 have the catalytic histidine and, at two residues C-terminal, the glutamate (Fig. 3), implying its human C3-like substrate specificity rather than human C4A or C4B-like. It is intriguing to note that fugu C4, which belongs to the carp C4-2 clade in the phylogenetic tree (Fig. 4), has a carp C4-1-like amino acid substitution, an asparagine in place of histidine, as in many of α2M homologues. This implies a functional importance of C4 molecule that lacks the catalytic histidine considering that this appears to be the only form of C4 present in fugu.

The x-ray crystallographic structures of human C3d and C4d have depicted a surface loop unique to C4d, spanning Ser1123 to Pro1236, which contains an acceptor site (Ser1217) for the covalent binding of C3b to form the classical pathway C5 convertase complex (C4bC2aC3b) (49). As seen in Fig. 3, the corresponding region of C4 from the lower vertebrates is much shorter (carp C4-1 and C4-2, shark C4, and medaka C4) than that of human C4 or even totally missing (Xenopus C4; data not shown). Although carp C4-1 and C4-2 have at least one serine or threonine in this region, carp C4-1 and C4-2 probably differ in the substrate specificity of the binding reaction of the thioester, i.e., C4-1 favors amino groups and C4-2 prefers hydroxyl groups. In addition to the catalytic histidine, a glutamate at its two residues C-terminal has been predicted to sharpen the substrate specificity. In human C3, the glutamate residue has been inferred to render the catalytic histidine with stronger nucleophilicity, providing a strict substrate specificity of the thioester toward hydroxyl-groups (48). Having a serine at the corresponding position, human C4B is reactive to both hydroxyl- and amino-groups (3, 49). Carp C4-2 and medaka C4 have the catalytic histidine and, at two residues C-terminal, the glutamate (Fig. 3), implying its human C3-like substrate specificity rather than human C4A or C4B-like. It is intriguing to note that fugu C4, which belongs to the carp C4-2 clade in the phylogenetic tree (Fig. 4), has a carp C4-1-like amino acid substitution, an asparagine in place of histidine, as in many of α2M homologues. This implies a functional importance of C4 molecule that lacks the catalytic histidine considering that this appears to be the only form of C4 present in fugu.
Effects of synthetic peptides representing the putative C2 binding site of C4-1 and C4-2 on the hemolytic activities of carp serum triggered via the classical and alternative pathways. A. Comparison of the putative C2 binding site sequence among human C4, carp C4-1 and carp C4-2. The corresponding region in the carp C3-H1 isoform is also included. The sequence stretch shown in bold in human C4 has been mapped as the primary C2 binding site. Sequences of the synthetic peptide with amidated C-termini are also shown in bold in the C4-1, C4-2 and C3-H1 of carp. The C1s cleavage site in C4 and the C3 convertase cleavage site in C3 are shown by a blank, B and C. Effect of the peptides on the classical pathway (B) and the alternative pathway (C) activities. RA was calculated from Y0 (percent hemolysis in the absence of the peptides) and Y (percent hemolysis in the presence of the peptide at the indicated concentration). Control reactions containing 10 mM EDTA showed no hemolysis for both target cells (data not shown).

C3 is yet to be determined. The observation that the C3-peptide inhibited both classical and alternative pathway-mediated hemolysis may imply functional importance of the amplification of the alternative pathway C3 activation for hemolytic reaction. The incomplete inhibition by the C3 peptide could be due to the relative abundance of C3 in carp serum, compared with C4 content. It is a little surprising that the two peptides gave a similar IC50 of ~60 μM despite their substantial sequence diversity. It is yet to be determined whether the two peptides compete for a common site in a single component of carp complement that is functionally equivalent to mammalian C2. In this context, the present result that neither C4 peptides inhibited the alternative pathway-initiated hemolysis is in disagreement with the idea that a single B/C2-like component functions both in the classical and alternative pathways in teleosts (51). Detailed analyses of the molecular composition and substrate specificities of the C3 convertases are needed to clarify the activation mechanism of the classical pathway in the bony fish complement.

In conclusion, the two divergent C4 isoatypes with and without the catalytic histidine probably have an ancient origin, which might have arisen before the divergence of bony and cartilaginous fish, playing a significant role in the classical/lectin pathway of the complement activation. Their functional diversity in the host defense is of great interest to better understand evolution of the complement system.

Acknowledgments
We are grateful to Dr. Kensuke Furukawa, Department of Bioscience and Biotechnology, Kyushu University, for N-terminal amino acid sequencing and to Satoru Totsuka and Takayuki Kajiya for technical assistance. The Boxshade server (www.ch.embnet.org/software/BOX_form.html) was used to format the multiple alignment.

Disclosures
The authors have no financial conflict of interest.

References


