PERGAMON

Developmental and Comparative Immunology 25 (2001) 419–430

Developmental & Comparative Immunology

www.elsevier.com/locate/devcompimm

Cloning and purification of the rainbow trout fifth component of complement (C5)

Silvia Franchini[a], Ioannis K. Zarkadis[b], Georgia Sfyroera[b], Arvind Sahu[a,1], William T. Moore[a], Dimitrios Mastellos[a], Scott E. LaPatra[c], John D. Lambris[a,*]

[a]*Department of Pathology and Laboratory Medicine, University of Pennsylvania, 401 Stellar-Chance Laboratories, Philadelphia, PA 19104, USA*

[b]*Department of Biology, School of Medicine, University of Patras 26500 Rion, Patra, Greece*

[c]*Clear Springs Foods Inc., Buhl, ID, USA*

Received 27 November 2000; received in revised form 17 January 2001; accepted 30 January 2001

Abstract

To gain further insight into the evolutionary history of the complement proteins C3, C4, and C5 we have now cloned the fifth component of complement from a rainbow trout (*Oncorhynchus mykiss*) liver cDNA library; this is the first report of C5 cloning in a species other than human and mouse. The deduced amino acid sequence of a partial cDNA clone (2.25 kb), representing approximately 44% of the coding sequence, showed 60 and 58% similarity to human and mouse C5, respectively. To validate the molecular information derived from the cloning we developed an improved purification protocol. Mass spectrometric analysis of C5 tryptic digests yielded peptide signals that matched theoretical protein sequence derived from the partial cDNA. Northern blot analysis of RNA from various tissues showed the presence of a single mRNA transcript in trout liver and Southern blot analysis indicated that the gene coding for C5 is present as a single copy in the trout genome. The presence of C5 in trout suggests that C3, C4, and C5 must have diverged before the appearance of teleost fish.

Keywords: Cloning; Teleost fish; Complement; C5; Mass spectrometry

1. Introduction

The complement system consists of a group of plasma proteins and cell receptors that play a critical role in host defense by interacting with components of both the innate and adaptive immunity [1,2]. Complement can be activated through three distinct pathways: the classical pathway, which is triggered by antigen-antibody complexes, and the alternative and lectin pathways, which are activated by direct binding of complement components to microbial surfaces [3]. Complement activation proceeds through a cascade of proteolytic cleavages that result in the formation of the membrane attack complex (MAC), or C5b-9 complex, which is responsible for the lysis of complement-targeted cells [4].

Activation of the fifth component of complement, C5, initiates the assembly of the MAC. Cleavage of C5 by either the classical (C4b,2a,3b) or the alternative (C3b, Bb; $C3b_2Bb$) pathway C5 convertase generates

* Corresponding author. Tel.: +1-215-746-5765; fax: +1-215-573-8738.

E-mail address: lambris@mail.med.upenn.edu (J.D. Lambris).

[1] A.S. is a Welcome Trust Overseas Senior Research Fellow in Biomedical Science at the National Centre for Cell Science, Pune University Campus, Ganeshkhind, Pune 411007, India.

0145-305X/01/$ - see front matter
PII: S0145-305X(01)00010-6

two proteolytic fragments, C5a and C5b. The newly generated C5b exposes a binding site for C6, leading to the assembly of a stable bimolecular complex, C5b-6. Sequential binding of C7, C8, and C9 to this original complex results in the formation of a transmembrane porous structure that can insert into lipid membranes and cause the direct lysis of cells. The anaphylatoxic fragment C5a participates in several inflammatory reactions exerting a potent chemoattractive effect on various myeloid cells including neutrophils, eosinophils, basophils, monocytes and mast cells [5]. C5a exerts its function by binding to the G-protein coupled receptor C5aR which is expressed on the surface of myeloid and, as recently described, non-myeloid cells [6].

Human C5 is present in plasma at a concentration of 75 μg/ml [7]. It is synthesized mainly in hepatocytes as an intracellular single-chain precursor with a βα-chain orientation and it is secreted as a two-chain glycoprotein (α and β, M_r 115,000 and 75,000) linked by disulfide bonds [8,9] . It is noteworthy, however, that C5 lacks the internal β-cysteinyl-γ-glutamyl thiol ester bond that is a common feature of C3, C4, and α2-macroglobulin [10].

C3, C4, and C5 are thought to have originated from a common ancestor by gene duplication [11–13]. This belief is based on the sequence similarity and exon-intron organization of these proteins [14]. It is, however, not known at what point in evolution the first gene duplication took place. Thus, probing the molecular structure and function of these proteins at various levels on the phylogenetic tree should lead to a detailed understanding of the origin of these proteins and of the complement system in general. To date, the complete amino acid sequence of C5 has been reported for only two higher vertebrates: human [15–17] and mouse [18–19]. Thus, in order to gain further insight into the functions of C5 and to elucidate its phylogenetic relationships with C3 and C4, we have cloned a gene encoding C5 from rainbow trout, a teleost fish representing a critical point in the evolutionary divergence of the complement system. In addition, we have developed an improved purification method that yields homogeneous C5 in a single chromatographic step.

In presenting the cloning of C5 from a lower vertebrate species we now demonstrate the structural conservation of this critical component of the lytic pathway of complement in phylogenetically divergent organisms, teleost fish and mammals.

2. Materials and methods

2.1. Chemicals

Chemicals for automated sequencing were obtained from Applied Biosystems (Foster City, CA). Restriction enzymes and trypsin were purchased from Promega (Madison, WI) and New England Biolabs, Inc. The nylon membranes used for hybridization were obtained from BioRad (Hercules, CA). All other chemicals and reagents used were reagent or higher grade.

2.2. Animals

Blood was collected from the caudal artery of rainbow trout (*Oncorhynchus mykiss*) (Clear Springs Foods, Inc. trout farm, ID) and immediately transferred to capped polypropylene 50-ml tubes containing 20 mM sodium citrate buffer on ice. Trout plasma was obtained by sequential centrifugation at 2000 *g* for 10 min at 4°C and 15,000 *g* for 20 min then stored at −70°C until processed further.

2.3. RNA isolation and cDNA library construction

RNA isolation and library construction were performed as previously described [20]. Fresh trout liver from a single animal was homogenized, and total RNA was extracted by the guanidine isothiocyanate method after precipitation with CsCl. Polyadenylated RNA was purified using the polyA Tract mRNA isolation kit (Promega, Madison, WI) according to the manufacturer's instructions. Double-stranded cDNA was synthesized from 10 μg of poly (A^+) RNA using a Time Saver cDNA synthesis kit (Amersham Pharmacia). After treatment with *EcoRI* methylase and addition of synthetic *EcoRI* linkers, the cDNA was digested with *EcoRI* and size-fractionated on a Sepharose CL4B column. cDNAs larger than 500 bp were ligated into *EcoRI*-digested, dephosphorylated λgt11 arms and packaged in vitro using the Amersham cloning system. The Y1090 strain of *Escherichia coli* was used as a host for infection by recombinant λ-phages.

2.4. Cloning of trout C5

During the trout C3 purification we identified and isolated a molecule with a two-chain structure similar in molecular weight to that previously published for trout C5 [21]. N-terminal sequencing of the molecule's α-chain showed high homology with human C5. The molecule's N-terminal peptide corresponding to the sequence SYDQQK was then used to design a degenerate oligonucleotide 5′-YTCYTTYT-GYTGRTCRTA-3′ (extends upstream). On the basis of the conserved amino acid sequence VFHLAG of human and mouse C5, a second oligonucleotide, 5′-GTNTTYCAYYTNGCNGGN-3′ (extends downstream), was also designed. Mixtures of nucleotides are represented by N = A, G, C and T; Y = C and T; R = A and G. These primers were subsequently used in RT-PCR reactions. Thirty cycles of amplification were conducted in a PCR Sprint thermocycler using the following parameters: 95°C for 0.5 min, 46°C for 0.5 min and 72°C for 1 min. The PCR product of the expected size (127 bp) was gel-purified and cloned into the p-GEMTeasy vector (Promega).

To isolate the trout C5 cDNA, a trout liver cDNA library containing 1.8×10^5 phages was screened under high stringency conditions (65°C) with a ^{32}P-labeled cDNA probe corresponding to the PCR product described above. The probe was labeled using the random primed DNA labeling kit (Amersham Pharmacia). Positive phages were plaque-purified and subcloned into the pIBI31 (IBI, New Haven, CTT) plasmid vector.

DNA sequence analysis of the trout C5 recombinant clones was performed by the dideoxy chain termination method [22] using the VISTRA 725 sequencer. Each sequence was determined at least twice from both strands.

2.5. Peptide synthesis and antibody production

A 14-amino acid peptide spanning the C-terminus of the trout C5 β-chain and the N-terminus of the α-chain (human C5 amino acid positions: 669–683) was synthesized using an Applied Biosystem 430A peptide synthesizer (Foster City, CA) as previously described [23]. The peptide was then coupled to BSA by the glutaraldehyde method and used for the immunization of rabbits and the production of immunoreactive antisera. The anti-peptide-specific antibody was purified by affinity chromatography using the synthetic peptide coupled to cyanogen bromide-activated Sepharose 4B (Pharmacia, Piscataway, NJ).

2.6. Purification and characterization of trout C5

Trout C5 was purified from plasma using a single chromatographic step. In brief, trout plasma (1 ml) was precipitated with 4% (w/v) polyethylene glycol (PEG) under constant stirring at 4°C for 30 min in the presence of the following inhibitors: 10 mM EDTA, 10 mM benzamidine, 0.001% NaN_3 and 1 mM phenylmethylsulfonyl fluoride (PMSF). The supernatant collected after centrifugation (10,000 *g* for 20 min at 4°C) was further precipitated with 16% PEG. After centrifugation the pellet was resuspended in 20 mM Na_2HPO_4 buffer, pH 7.8, and loaded onto a Mono Q HR5/5 anion exchange chromatography column (Pharmacia) pre-equilibrated with the same buffer. Trout C5 was eluted with a linear gradient of NaCl (0–250 mM), and C5-containing fractions were identified by SDS–polyacrylamide gel electrophoresis and immunoblotting using the affinity-purified trout C5 anti-peptide antibody. The immunoreactive fractions were pooled and dialyzed overnight against phosphate-buffered saline, pH 7.4, at 4°C. The purity of the C5 preparation exceeded 95%, as determined by SDS–PAGE analysis.

The serum-purified protein was characterized by mass spectrometric analysis and Edman sequencing.

2.6.1. Tryptic peptide mass mapping by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS).

Tryptic peptides derived from the trout C5 chains (~30 pmoles each of the α- and β-chains) were fractionated by SDS–PAGE at 7.5% acrylamide concentration under reducing conditions (inset, Fig. 6). Gel slices were subjected to an in-gel tryptic digestion procedure: [24]. Gel slices were equilibrated with 100 mM ammonium bicarbonate buffer, dehydrated, and re-swelled in buffer containing ~2 pmole of sequencing grade porcine trypsin (E/S = ~1/15). Tryptic peptides extractable with 60% acetonitrile/0.1% TFA were subjected to MALDI-MS analysis using a Micromass (Beverly, MA) TofSpec 2E time-of-flight mass spectrometer (1.0 m flight tube)

equipped with a nitrogen laser (337 nm) and a reflectron. Data were collected in the positive ion mode using the reflectron. Accelerating and detector voltages were set at 20 and 1.75 kV, respectively. Peptide samples were desalted by a matrix para-crystalline method previously described [25]. The mass spectrometer was calibrated using synthetic peptide external standards, resulting in a mass measurement accuracy within 0.3 amu below mass 4500. Mass matching and peptide identification was assisted by mass-matching BIOLNYX algorithms provided by the manufacturer's (Micromass, Beverly, MA) software package MASSLYNX.

2.6.2. Amino terminal sequence analysis by automated Edman degradation chemistry.

Amino acid sequencing of electroblotted proteins was accomplished using a PE-ABI Model 473A Protein Sequencer (Foster City, CA) equipped with on-line PTH-amino acid analysis, using programmed chemistry cycles and HPLC operation programs provided by the manufacturer. Amino acid determinations were made by visual inspection of the data. An internal sequencing standard, a 17-residue synthetic peptide containing the atypical residues norleucine and succinyllysine, was used to monitor sequencer performance and to establish whether a protein is N-terminally modified blocking the initiation of the Edman chemistry

2.7. Northern blot analysis

To analyze the tissue distribution of trout C5, tissue samples including brain, intestine, kidney, and liver were freshly harvested and immediately frozen in liquid nitrogen. Twelve micrograms of total RNA from each tissue were denatured using formamide, separated by electrophoresis on a 1% agarose gel, and blotted onto a nylon membrane (Zeta-Probe membrane, Biorad) as previously described [26]. Prehybridization was performed for 30 min at 43°C in a solution containing 50% formamide, 0.12 M sodium phosphate (pH 7.2), 0.25 M sodium chloride, 7% SDS and 1 mM EDTA. The membrane was hybridized at 43°C for 16 h with the ^{32}P-radiolabeled probe of 700 bp representing the 5′ end of the partial cDNA clone (2226 bp) obtained from the screening of the library (digestion with *EcoRI* and *AccI*). The membrane was washed twice for 15 min at room temperature in 2× standard sodium citrate and 0.1% SDS buffers. The film was developed after 16 h of exposure.

2.8. Southern blot analysis

To determine whether C5 is present as single or multiple copies within the trout genome, Southern blot analysis was performed using the same probe described for the northern blot. In brief, genomic DNA was isolated from trout erythrocytes using the phenol-chloroform method, and 12 μg of DNA were digested to completion with 10 units/μg of restriction endonucleases (*EcoRI*, *HindIII*, *PstI*). The digests were subsequently electrophoresed on a 0.8% agarose gel and transferred to a nylon membrane. After incubation at 80°C for 2 h, the membrane was prehybridized for 30 min at 65°C in a solution containing 7% SDS, 0.5 M sodium phosphate (pH 7.2) and 0.1 M EDTA. The membrane was then hybridized with the ^{32}P-labeled probe at 65°C for 16 h and washed twice for 20 min with 40 mM Na_2HPO_4 and 1 mM EDTA in 5% SDS at room temperature. The X-ray film (Kodak) was developed after 2 days of exposure.

2.9. Computer methods

Multiple alignments were generated using the Clustal W 1,5 program [27] and the obtained alignment was used to construct a phylogenetic tree using the neighbor-joining method (MEGA) [28].

3. Results

3.1. Isolation and sequence analysis of a partial cDNA encoding trout C5

To clone the gene encoding C5 in trout, we generated a cDNA probe by RT-PCR using trout liver mRNA and degenerate primers as shown in Fig. 1. The primers were designed on the basis of the following protein sequences: (1) -VFHLAG-, which is found to be conserved in human and mouse C5, and (2) -SYDQQK-, which was obtained by N-terminal sequencing of trout C5. RT-PCR amplification resulted in a single DNA band of the expected size (126 bp), which corresponded to positions 1935–2061 of human C5 cDNA. The DNA was gel-purified and

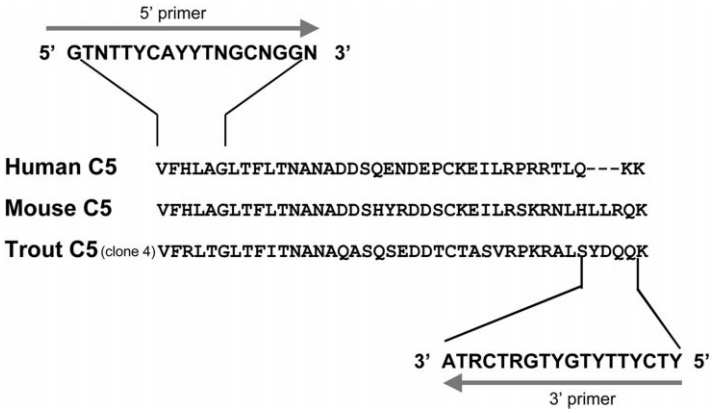

Fig. 1. RT-PCR amplification scheme of a trout C5 cDNA fragment. The primers used in the RT-PCR reaction were designed based on the amino acid sequence (-VFHLAG-) which is highly conserved among human and mouse C5. The anti-sense primer was designed based on the trout N-terminal sequence (-SYDQQK).

subcloned into the p-GEM^{T}easy vector (Promega). Then, the nucleotide sequence of the insert was determined, and its deduced amino acid sequence showed 40, 27 and 29% identity with the corresponding regions of human C5, C4, and C3, respectively. Therefore, this cDNA clone (clone 4) was identified as trout C5. Subsequently, the PCR-amplified clone 4 was used as a probe to screen a trout λgt11 library under high-stringency conditions (65°C). Among the 18×10^4 clones, approximately 40 positive signals were detected, and 12 clones were plaque-purified and analyzed by restriction enzyme digestion using *EcoR1*. Subsequently, the nucleotide sequence was determined. The longest clone was 2226 bp and predicted a single open reading frame of 745 amino acids corresponding to residues 388–1119 of human C5.

3.2. Alignment of the trout C5 amino acid sequence with the C3, C4, C5 and α2M sequences from various species

The deduced amino acid sequence of trout C5 was aligned with those of C5, C4, and C3 from representative species using Clustal W software. Part of this alignment containing only human and mouse C5 is shown in Fig. 2. The amino acid identity of trout C5 was 40% with human C5 and 39% with mouse C5. Lower degrees of amino acid identity, of 27, 29 and 23%, were observed between the trout C5 amino acid sequence and the human C4, C3, and α2M sequences, respectively.

Based on the similarity of trout C5 with various members of the α2M protein family, a phylogenetic tree was constructed using the neighbor-joining method. Trout C5 clustered with human and mouse C5 (Fig. 3).

3.3. Isolation and characterization of the trout C5 protein

To isolate the protein coded by the gene related to the clone under study we developed an improved two-step purification method. Plasma was fractionated by PEG precipitation and then subjected to anion-exchange chromatography. In order to detect the protein of interest throughout the purification procedure we generated an antibody against a peptide corresponding to residues 669–683 of human C5, spanning the C-terminus of the C5 β-chain and the N-terminus of the α-chain.

This antibody reacted in western blot analysis only with the β-chain of trout C5, indicating that after the coupling to BSA either the C-terminus portion was only exposed or that the N-terminus portion is not antigenic. Homogeneous trout C5 was obtained after a single chromatographic step, as judged by Coomasie blue staining of the eluted fractions under reducing conditions (Fig. 4, left panel).

Immunoblotting of the eluted fractions after anion exchange chromatography showed reactivity only in a single region of the elution profile. Furthermore, western blot analysis of total serum and of fractions resulting from various steps of the PEG precipitation revealed again the presence of a single band.

```
TC5    FIKNKLGVDEKLQCHTLTVDNSKRTKDDGIALFICNLHSEAESASFIFKTADERL    55
MC5    VTLMAQTVDVNQETSDLETKRSITHDTDGVAVFVLNLPSNVTVLKFEIRTDDPEL    55
HC5    VILNAQTIDVNQETSDLDPSKSVTRVDDGVASFVLNLPSGVTVLEFNVKTDAPDL    55
        .    ..* . .  .*. ..*    .**.* *. **.*...  .*  .*..  *

TC5    PFGSQAEYTLEAKAYSSPNERYLYIDLPSDYSGLQVGQHVSIQIYFHARSYL-RI   109
MC5    PEENQASKEYEAVAYSSLSQSYIYIAWTENYKPMLVGEYLNIMVT-PKSPYIDKI   109
HC5    PEENQAREGYRAIAYSSLSQSYLYIDWTDNHKALLVGEHLNIIVT-PKSPYIDKI   109
       * ..**  .  * **** ...*.**. ..... . **....* .  . ..*. .*

TC5    ETFSYQIISRGKIVKFASEKRLQGTGSQAISFTVTSDMVPSIRLLVYYILHGEKT   164
MC5    THYNYLILSKGKIVQYGTREKLFSSTYQNINIPVTQNMVPSARLLVYYIVTGEQT   164
HC5    THYNYLILSKGKIIHFGTREKFSDASYQSINIPVTQNMVPSSRLLVYYIVTGEQT   164
       . ..* *.*.***..... ... ... *.*...** .**** *******. **.*

TC5    PELVADSVWIDVKDKCVNALKTELSFRKKDYKPKDELIFEVKTGQDSLVALSAID   219
MC5    AELVADAVWINIEEKCGNQLQVHLSPDEYVYSPGQTVSLDMVTEADSWVALSAVD   219
HC5    AELVSDSVWLNIEEKCGNQLQVHLSPDADAYSPGQTVSLNMATGMDSWVALAAVD   219
       .***.*.**.....** *.*...**     *.* ... ... *. ** ***.*.*

TC5    TAVYSLRAPLQDPVTRVLRHIEQ-SDQGCGGGGGKDNADVFRLTGLTFITNANAQ   273
MC5    RAVYKVQGNAKRAMQRVFQALDEKSDLGCGAGGGHDNADVFHLAGLTFLTNANAD   274
HC5    SAVYGVQRGAKKPLERVFQFL-EKSDLGCGAGGGLNNANVFHLAGLTFLTNANAD   273
        *** ..   . .. **.. . . ** ***.*** .**.**.*.****.*****.

       β-chain ⟵ linker ⟶ α-chain
TC5    ASQSEDDTCTASVRPKRALSYDQQK--EKAKQYGKDFF--CCWQGLHQIPTLDTC   324
MC5    DSHYRDDSCKEILRSKRNLHLLRQKIEEQAAKYKHSVPKKCCYDGAR-VNFYETC   328
HC5    DSQENDEPCKEILRPRRTLQ---KKIEEIAAKYKHSVVKKCCYDGAC-VNNDETC   324
       .*.  *..*.. .*..*.*    .*  * * .* .    **.*       **
                                  C5 convertase cleavage site
TC5    AHRANRLKVKLPSDVKKCRAVFTECCDYSLQLLQTSSDHV--LARVEMAILFDLM   377
MC5    EERVARVTI-----GPLCIRAFNECCTIANKIRKESPHKPVQLGRIHIKTLLPVM   378
HC5    EQRAARISL-----GPRCIKAFTECCVVASQLRANISHKDMQLGRLHMKTLLPVS   374
       ..*..*...        *  .*.***  . ..  . ...   *.*... .*. .

TC5    PTQIRKLTFPESWLWEVQPVRSGQITLNRNLPDSLTTWEIKAVGVFKKEGSSGIC   432
MC5    KADIRSY-FPESWLWEIHRVPK-RKQLQVTLPDSLTTWEIQGIGI----SDNGIC   427
HC5    KPEIRSY-FPESWLWEVHLVPR-RKQLQFALPDSLTTWEIQGIGI----SNTGIC   423
        ..**.  ********.. *.. .  *. .**********...*.    ...***

TC5    VADPIKVSVTQAVSVDVPLPYSMVRGEQIELRGSVYNQAEDNIKYCVTLTAGAGV   487
MC5    VADTLKAKVFKEVFLEMNIPYSVVRGEQIQLKGTVYNYMTSGTKFCVKMSAVEGI   482
HC5    VADTVKAKVFKDVFLEMNIPYSVVRGEQIQLKGTVYNYRTSGMQFCVKMSAVEGI   478
       ***..*..* ..* ... .***.******.*.*.***  ... ..**...* .*.

TC5    CLFQGKPTKDDGIKTTNCDKTSLLESGSVGLVTFTLMALEVGSHTLTFTLHTDKS   542
MC5    CTSGSSAASLHTSRPSRCVFQRI-EGSSSHLVTFTLLPLEIGLHSINFSLET---   533
HC5    CTSESPVIDHQGTKSSKCVRQKV-EGSSSHLVTFTVLPLEIGLHNINFSLET---   529
       *   .   . .. ....*   .. *..*  *****...**.* *...*.*.*

TC5    RFNDQI-VKTIRVVPEGIRTEDLSGGKLDPQGLYGAASRKVELSNTIPSQLVPKT   596
MC5    SFGKDILVKTLRVVPEGVKRESYAGVILDPKGIRGIVNRRKEFPYRIPLDLVPKT   588
HC5    WFGKEILVKTLRVVPEGVKRESYSGVTLDPRGIYGTISRRKEFPYRIPLDLVPKT   584
        *...* ***.******.. *. .*  ***.*. *  .*. *..  ** .*****

TC5    SVERLLTINGEILGEVLAILNNPDGLRQLVNLPSGSAEGEIMRVLPIYFVYHYLE   651
MC5    KVERILSVKGLLVGEFLSTVLSKEGINILTHLPKGSAEAELMSIAPVFYVFHYLE   643
HC5    EIKRILSVKGLLVGEILSAVLSQEGINILTHLPKGSAEAELMSVVPVFYVFHYLE   639
       ...*.*...* ..** *. . . .*.. *..**.****.*.*.. *...*.****

TC5    TTGRWDIMGLEEHNSPLELKRKIREGITSILSFKRPREFSYSMWKDKEASTWLTA   706
MC5    AGNHWNIFYPDTLSKRQSLEKKIKQGVVSVMSYRNA-DYSYSMWKGASASTWLTA   697
HC5    TGNHWNIFHSDPLIEKQKLKKKLKEGMLSIMSYRNA-DYSYSVWKGGSASTWLTA   693
       ....*.*.  .   .  .*..*...*. *..*.... ..***.**. .*******

TC5    LVVKTLGQVDKYVKVDSDMLSNSIFWLINKAQNEDGSFR   745
MC5    FALRVLGQVAKYVKQDENSICNSLLWLVEKCQLENGSFK   736
HC5    FALRVLGQVNKYVEQNQNSICNSLLWLVENYQLDNGSFK   732
       .....****.***. . . ..**..**... * ..***.
```

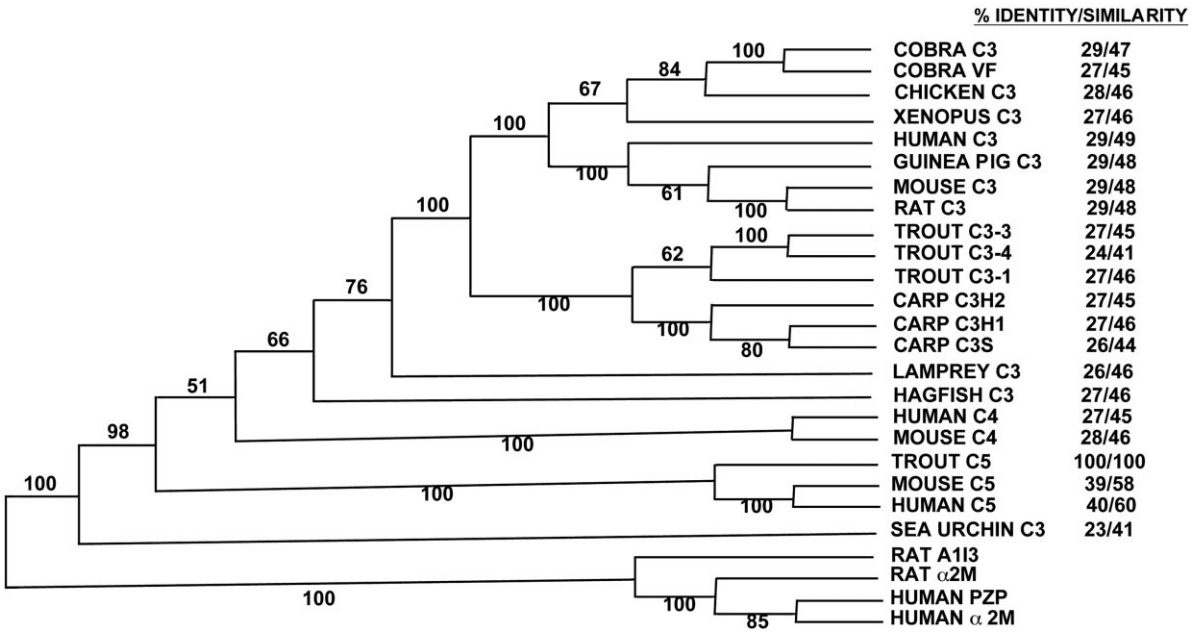

Fig. 3. Phylogenetic tree showing the relation of C5 to other thioester-containing proteins. The relationships among human C3, human C4, human C5, human PZP, human α2M, mouse C3, mouse C4, mouse C5, rat C3, rat A1I3, rat α2M, guinea pig C3, chicken C3, cobra C3, cobra VF, trout C3-3, trout C3-4, trout C3-1, carp C3H2, carp C3H1, carp C3S, xenopus C3, lamprey C3, hagfish C3, and trout C5 was analyzed by the neighbor-joining method (MEGA) for their entire amino acid sequences, based on the alignment performed with CLUSTAL W software. The numbers on the branches show the percent recovery in 1000 bootstrap replications. The numbers on the right indicate the percent amino acid sequence identity and similarity among trout C5 and C5, C3, and C4 of various species.

Edman degradation of the C5 α-chain generated an amino acid sequence identical to that deduced from the cDNA clone (Fig. 5, underscored residue region). The negative sequencing result obtained for an equimolar amount of β-chain and the positive co-sequencing run obtained on the internal standard for the sequencing run suggest that the N-terminal amino acid of the β-chain was not available for Edman degradation. That the β-region band did indeed represent the β-chain was confirmed by mass spectrometric analysis of the tryptic peptides derived from this band and by mass-matching to theoretical tryptic peptide sequences that were derived from the cDNA clone isolated in this study.

Prominent mass spectral signals were obtained for tryptic peptides that cover or account for 31% of the known cDNA-derived amino acid sequence. The ion signals that had masses consistent with those calculated from sequence derived from cDNA sequence are indicated by an asterisk in the MALDI-mass spectrum shown in Panel A of Fig. 5. The amino acid sequences represented by tryptic peptide ion signal masses are shown in bold face in Panel B of Fig. 5.

3.4. Northern and Southern blot analyses

To examine the tissue distribution profile of trout C5 mRNA, northern blot analysis was performed using total RNA from several trout tissues. The ^{32}P-labeled cDNA probe detected a single, 5-kb band only in trout liver (Fig. 6); brain, intestine, kidney showed no hybridization, suggesting that trout C5 is expressed

Fig. 2. Alignment of the deduced amino acid sequence of trout C5 with human and mouse C5. The alignment was performed using the Clustal W 1.5 program. Dots represent positions in the aligned sequence that are well conserved; stars indicate amino acids identical in all three sequences, and dashes show gaps introduced to increase the identity. Boxed sequence corresponds to the α–β junction. The amino acid numbers are given on the right.

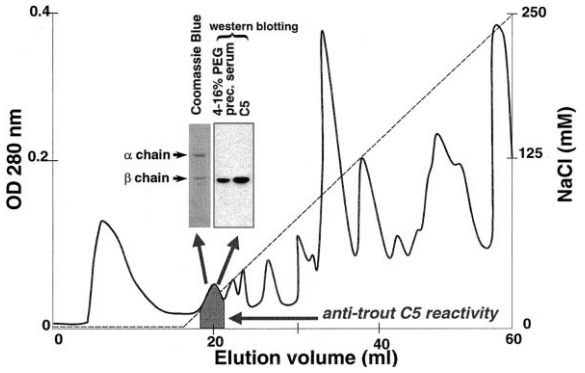

Fig. 4. Elution profile, SDS–PAGE and Western blot analysis of trout C5. 4–16% PEG-precipitated trout serum was injected onto a MonoQ HR5/5 column equilibrated with 20 mM Na_2HPO_4, pH 7.8. C5 was eluted with a 20 ml gradient of NaCl (0–500 mM) at a flow rate of 1 ml/min. —, $A_{280;}$ – – –, concentration of NaCl. The shaded peak represents the fractions containing C5. The right panel shows the reactivity of these fractions with an anti-trout C5 antibody that recognizes the β-chain. The left panel represents an SDS–PAGE pattern of homogeneous trout C5 (3 μg) resolved under reducing conditions and stained with Coomasie blue.

only in the liver. In order to estimate the copy number of this gene, we performed Southern blot analysis using the same probe used for the northern blot. Only a single hybridizing band was detected when a panel of restriction enzymes (EcoRI, HindIII and PstI) was employed, suggesting that the trout C5 gene is present as a single copy in the trout genome (Fig. 7).

4. Discussion

C5 represents a critical component of the complement system, since it is the initial molecule in the assembly of the cytolytic MAC and a key mediator in several inflammatory processes. Various components of the lytic pathway (MAC) have been identified in two teleost fish, trout [29] and carp [30]. They have been found to have similar chain structure to the mammalian MAC.

Several studies [31,32] have raised the question whether C5 predated C3 and C4 in the divergence

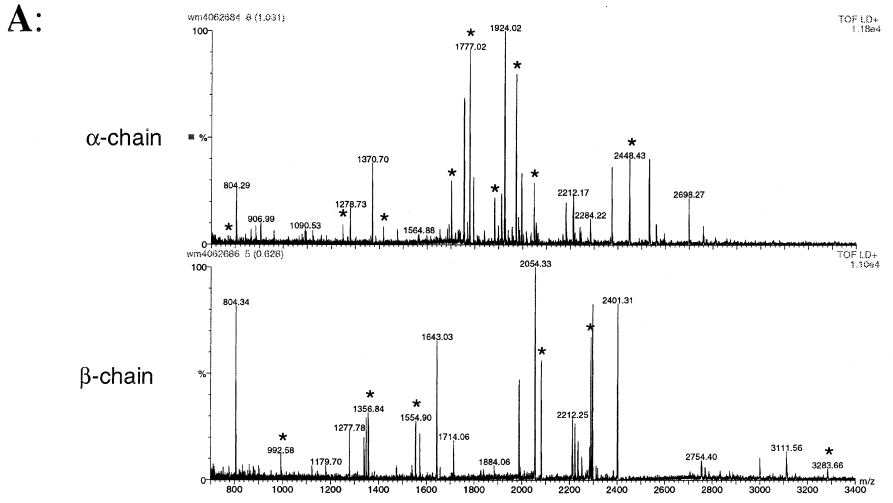

Fig. 5. Mass spectrometric analysis (MALDI) of trout C5. Panel A shows the spectra corresponding to the tryptic digests of the α- and β-chains, respectively. Panel B represents the deduced amino acid sequence of trout C5 in which tryptic peptides identified by mass spectrometry (see asterisks in panel B) are highlighted. The underlined sequence indicates the peptide identified by Edman degradation.

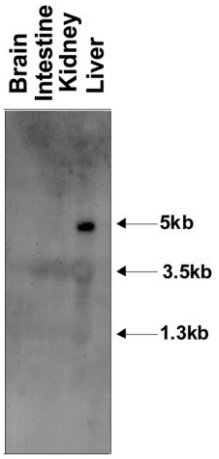

Fig. 6. Northern blot analysis of trout C5. Twelve micrograms of total RNA from various trout tissues were denatured with formamide, separated by electrophoresis on a 1% agarose gel, and blotted onto a nylon membrane. After hybridization a band of 5 kb was observed in the line representing total RNA from liver. The 28 and 18 S rRNA subunits served as internal molecular weight markers.

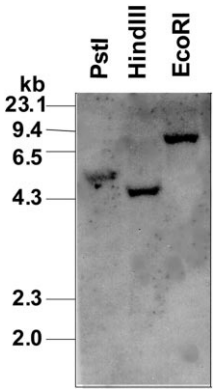

Fig. 7. Southern blot analysis of trout C5. Twelve micrograms of trout genomic DNA were digested with the restriction enzymes *EcoRI*, *HindIII* and *PstI*, separated by electrophoresis on a 0.8% agarose gel, and blotted onto a nylon membrane. Hybridization with a DNA probe was performed for 16 h at 65°C. The positions of DNA standards are shown at left.

of the three molecules from a common ancestral molecule. These issues still remain unresolved because of a lack of sufficient structural/functional information about C3, C4, and C5 in lower vertebrates. C3 has been extensively characterized in all major groups of vertebrate [33], including invertebrate species such as ascidians [34] and the sea urchin [35]. Furthermore, the primary structure of C4 has been reported in teleost fish [36] and amphibians [37]. However no data have been generated thus far concerning the structure of C5 in lower vertebrates and to date, characterization of this molecule has been reported only for humans [15] and mice [18].

Therefore, to gain further insight into the evolutionary history of C5 and to study its phylogenetic relationships with C3 and C4 we have now cloned the gene encoding C5 in a teleost fish, the rainbow trout. In addition to facilitating future studies of the role of complement in immune recognition and helping to delineate the mechanism by which it exerts its functions, this study has produced an improved purification method that yields homogeneous C5 in a single chromatographic step, in contrast to the three chromatographic steps used in a previously published method [21].

In this study we report the amino acid sequence of a cDNA clone encoding for the C-terminal 285 amino acid residues of the C5 β-chain, the entire C5a fragment, and the N-terminal 379 residues of the α'-chain. Several lines of evidence indicate that the DNA sequence we obtained is that of trout C5: First, the DNA sequence shows high similarity to that of human (60%) and mouse C5 (58%) and very low similarity to that of C3 cloned from trout or other species ($<$46%). Second, it is of particular note that tryptic peptide mass mapping of the purified protein showed at least 31% coverage of the protein level information with that deduced from cDNA. Several dominant ion signals derived from both the α- and β-subunits indicate that the serum protein is the product of the gene represented by this partial cDNA clone. It should be noted that the mass spectral results were obtained on a total protein digest. Complete accounting of every possible tryptic peptide is not expected when a total in-gel digest is presented to the mass spectrometer as a single analysis. This is partly attributed to ion suppression effects that mask some peptide signals as well as to incomplete recovery of peptides in the in-gel digestion technique. Nevertheless, the high number of sequence/mass-assignable peptides demonstrates that the purified material is assuredly derived from a gene represented by the cDNA clone analyzed. Ion signals that are not assigned may represent post-translationally modified peptides or partial tryptics that cannot be unambiguously placed, given

mass data alone. A third line of evidence that the reported sequence represents indeed the C5 cDNA clone is the fact that the trout C5 transcript lacks an internal thiolester bond, as previously reported [38]. Examination of the deduced amino acid sequence of trout C5 indicates that the molecule is synthesized as a single-chain precursor with a β/α orientation and a conserved linker sequence -RPKR- at the junction between the two chains. It is interesting that the residue Arg-78 at the C5 convertase cleavage site of human C5 (Arg-Leu) is conserved in the trout whereas the Leu residue has been replaced by Val.

The deduced trout C5a peptide consists of 77 amino acids, with all six cysteine residues conserved as in mouse [18], and rat C5a [39]. Of note is the fact that trout C5a contains three additional amino acids at its amino terminus that are not present in the human [15], bovine [40], or porcine [41] C5a sequences. This feature is shared with mouse and rat C5a. Studies of the spasmogenic activity of rat C5a in guinea pig ileum contraction have reported a 100- to 1000-fold higher activity than that for human C5a, suggesting that these N-terminal amino acids are essential for the stabilization of the active conformation of the molecule [39].

A number of critical residues involved in the interaction with C5aR have been identified by site-directed mutagenesis studies of the human C5a molecule [42–44]. Results of these studies suggest that three discontinuous regions of the human C5a molecule, including residues Arg-74, Leu-72, and Lys-68, appear to act in concert to achieve binding specificity and full potency. Protein sequence alignment of the trout and human C5a carboxyl-terminal regions showed that Arg-74 and Leu-72 are conserved in the trout, whereas Lys-68 has been replaced by His. Furthermore, our analysis revealed that Arg-40, in the disulfide-linked core region, is also conserved. NMR studies have shown that substitution in this position of the core causes a localized conformational change that is associated with loss of potency. However, other investigators have reported that replacement of this core residue Arg-40 with Gly does not significantly affect functional C5a activity [44].

Ala-26, located in the region between the core and the amino terminal helix, is highly conserved in all species whose sequences have been reported thus far. However, in trout C5a this residue has been replaced by Leu. Furthermore, trout C5 includes a five-amino acid insertion at position Val-43 (of human C5a) that is not present in any other C5a that has been sequenced; the effect that these residues have on the biological activity of the peptide is currently not known.

The striking conservation of several functionally important residues that we have found in this lower vertebrate species may be useful in determining a minimal structure that can retain binding efficiency and biological potency.

Northern blot analysis revealed the presence of a single mRNA transcript of approximately 5kb in trout liver suggesting that C5 is synthesized in the liver as a single isoform. Furthermore, Southern blot analysis of trout genomic DNA suggested that the gene encoding trout C5 is present as a single copy in the trout genome.

The construction of a phylogenetic tree using the aligned protein sequences of several members of the α2M family, including the partial sequence of trout C5 reported in this study, revealed that trout C5 is grouped in a common cluster with human and mouse C5. This finding indicates that C5 has retained a high degree of structural conservation throughout evolution, which is consistent with its critical role in several immunoregulatory and effector pathways of the innate immune response.

Concerning the evolution of C3, C4, and C5 a prevailing hypothesis [14] is that these proteins were generated by gene duplication from α2M. One of the duplicated α2M genes must have been modified to form a C3/C4/C5 ancestral gene, which was then duplicated a second time to give rise to C3, C4 and C5. The exact points of these gene duplications, however, are not yet known. Preliminary results in our laboratory have indicated the presence of a C4 molecule in trout. This finding, together with the presence of C5 in trout, suggests that C5, C3, and C4 must have diverged before the appearance of teleost fish.

Acknowledgements

We thank L Spruce and Y. Harrison-Shahan for excellent technical assistance and Dr D. McClellan for editorial assistance. This work was supported by

National Institutes of Health grants AI 30040 and GM 56698 and Cancer Core Support Grant CA 16520 and by the Research Foundation, University of Patras (K. Karatheodori, Award). The sequence described in this paper has been deposited in the GenBank databas under accession number AF 349001.

References

[1] Lambris JD. The third component of complement: chemistry and biology, 1st ed. Berlin: Springer-Verlag, 1990.
[2] Holers MV. Complement receptors. Year in Immunology 1989;4:231–40.
[3] Holers MV. Complement. In: Rich RR, Fleischer TA, Shearer WT, Schwartz BD, Strober W, editors. Clinical immunology: principles and practice. St Louis: Mosby, 1996. p. 363–91.
[4] Lambris JD, Sahu A, Wetsel R. The chemistry and biology of C3,C4, and C5. In: Volanakis JE, Frank M, editors. The human complement system in health and disease. New York: Marcel Dekker Inc, 1998. p. 83–118.
[5] Morgan BP. Regulation of the complement membrane attack pathway. Crit Rev Immunol 1999;19(3):173–98.
[6] Zwirner J, Fayyazi A, Gotze O. Expression of the anaphylatoxin C5a receptor in non-myeloid cells. Mol Immunol 1999;36(13-14):877–84.
[7] Kohler PF, Müller-Eberhard HJ. Immunochemical quantitation of the third, fourth and fifth components of human complement: concentrations in the serum of healthy adults. J Immunol 1967;99:1211–6.
[8] Nilsson UR, Mandle RJ, McConnell-Mapes JA. Human C3 and C5: subunit structure and modification by trypsin and C42-C423. J Immunol 1975;114:815–22.
[9] Tack BF, Morris SC, Prahl JW. Fifth component of human complement: purification from plasma and polypeptide chain structure. Biochemistry 1979;18:1490–7.
[10] Sottrup-Jensen L, Stepanik TM, Kristensen T, Lonbald PB, Jones CM, Wierzbicki DM, Magnusson S, Domdey H, Wetsel RA, Lundwall A. Common evolutionary origin of alpha 2-macroglobulin and complement components C3 and C4. Proc Natl Acad Sci USA 1985;82:9–13.
[11] Dodds AW, Day AJ. The phylogeny and evolution of the complement system. Immunol and Med 1993;20:39–88.
[12] Moon KE, Gorski JP, Hugli TE. Complete primary structure of human C4a anaphylatoxin. J Biol Chem 1981;256(16):8685–92.
[13] Fernandez HN, Huglij TE. Chemical evidence for common genetic ancestry of complement components C3 and C5. J Biol Chem 1977;252(5):1826–8.
[14] Haviland DL, Haviland JC, Fleischer DT, Wetsel RA. Structure of the murine fifth complement component (C5) gene. A large, highly interrupted gene with a variant donor splice site and organizational homology with the third and fourth complement component genes. J Biol Chem 1991;266: 11818–25.
[15] Haviland DL, Haviland JC, Fleischer DT, Hunt A, Wetsel RA. Complete cDNA sequence of human complement pro-C5. Evidence of truncated transcripts derived from a single copy gene. J Immunol 1991;146:362–8.
[16] Lundwall AB, Wetsel RA, Kristensen T, Whitehead AS, Woods DE, Ogden RC, Colten HR, Tack BF. Isolation and sequence analysis of a cDNA clone encoding the fifth complement component. J Biol Chem 1985;260:2108–12.
[17] Wetsel RA, Lemons RS, Le Beau MM, Barnum SR, Noack D, Tack BF. Molecular analysis of human complement component C5: localization of the structural gene to chromosome 9. Biochemistry 1988;27:1474–82.
[18] Wetsel RA, Ogata RT, Tack BF. Primary structure of the fifth component of murine complement. Biochemistry 1987;26:737–43.
[19] Wetsel RA, Fleischer DT, Haviland DL. Deficiency of the murine fifth complement component (C5). A 2- base pair gene deletion in a 5′-exon. J Biol Chem 1990;265: 2435–40.
[20] Zarkadis IK, Sarrias MR, Sfyroera G, Sunyer JO, Lambris JD. Cloning and structure of three rainbow trout C3 molecules: a plausible explanation for their functional diversity. Dev Comp Immunol 2001;25(1):11–24.
[21] Nonaka M, Natsuume-Sakai S, Takahashi M. The complement system in rainbow trout (Salmo gairdneri) II. Purification and characterization of the fifth component (C5). J Immunol 1981;126(4):1495–8.
[22] Sanger FS, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 1977;74:5463.
[23] Becherer JD, Lambris JD. Identification of the C3b receptor-binding domain in third component of complement. J Biol Chem 1988;263:14586–91.
[24] Williams KR, et al. Internal protein sequencing of SDS PAGE-separated proteins: a collaborative ABRF study. In: Marshak DR, et al., editors. Techniques in protein chemistry, vol. 8. St Diego: Academic Press, 1997. p. 99–109.
[25] Fan X, Beavis RC. Growing protein-doped sinaptic acid crystals for laser-desorption. An alternative preparation method for difficult samples. Org Mass Spectrom 1993;28(12): 1424–9.
[26] Sunyer JO, Zarkadis I, Sarrias MR, Hansen JD, Lambris JD. Cloning, structure, and function of two rainbow trout Bf molecules. J Immunol 1998;161(8):4106–14.
[27] Thompson J, Higgins D, Gibson T. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673.
[28] Saitou N, Nei M. The neighbor-join method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–25.
[29] Nonaka M, Natsuume-Sakai S, Takahashi M. The complement system in rainbow trout (Salmo Gairdneri). J Immunol 1981;126:1495–8.
[30] Nakao M, Uemura T, Yano T. Terminal components of carp complement constituting a membrane attack complex. Mol Immunol 1996;33(11-12):933–7.
[31] Nonaka M, Takahashi M. Complete complementary DNA

sequence of the 3rd component of complement of Lamprey: implication for the evolution of thioester containing proteins. J Immunol 1992;148:3290–5.
[32] Hughes AL. Phylogeny of the C3/C4/C5 complement-component gene family indicates that C5 diverge first. Mol Biol Evol 1994;11:417–25.
[33] Sunyer JO, Lambris JD. Evolution and diversity of the complement system of poikilothermic vertebrates. Immunol Rev 1998;166:39–57.
[34] Ji X, Azumi K, Sasaki M, Nonaka M. Ancient origin of the complement lectin pathway revealed by molecular cloning of mannan binding protein-associated serine protease from a urochordate, the Japanese ascidian, Halocynthia roretzi. Proc Natl Acad Sci USA 1997;94(12):6340–5.
[35] Al-Sharif WZ, Sunyer JO, Lambris JD, Smith LC. Sea urchin coelomocytes specifically express a homoloque of the complement component C3. J Immunol 1997;160:2983–97.
[36] Kuroda N, Naruse K, Shima A, Nonaka M, Sasaki M. Molecular cloning and linkage analysis of complement C3 and C4 genes of the Japanese medaka fish. Immunogenetics 2000;51(2):117–28.
[37] Mo R, Kato Y, Nonaka M, Nakayama K, Takahashi M. Fourth component of *Xenopus laevis* complement: cDNA cloning and linkage analysis of the frog MHC. Immunogenetics 1996;43(6):360–9.
[38] Janatova J, Tack BF. Fourth component of human complement: studies of an amine-sensitive site comprised of a thiol component. Biochemistry 1981;20:2394–402.
[39] Cui L, Carney DF, Hugli TE. Primary structure and functional characterization of rat C5a: an anaphylatoxin with unusually high potency. Protein Sci 1994;3(8):1169–77.
[40] Gennaro R, Simonic T, Negri A, Mottola C, Secchi C, Ronchi S. C5a fragment of bovine complement. Purification, bioassays, amino-acid sequence and other structural studies. Eur J Biochem 1986;155:77–86.
[41] Gerard C, Hugli TE. Amino acid sequence of the anaphylatoxin from the fifth component of porcine complement. J Biol Chem 1980;255(10):4710–5.
[42] Mollison KW, Mandecki W, Zuiderweg ER, Fayer L, Fey TA, Krause RA, Conway RG, Miller L, Edalji RP, Shallcross MA. Identification of receptor-binding residues in the inflammatory complement protein C5a by site-directed mutagenesis. Proc Natl Acad Sci USA 1989;86(1):292–6.
[43] Mollison KW, Fey TA, Krause RA, Miller L, Edalji RP, Conway RG, Mandecki W, Shallcross MA, Kawai M, Or YS. C5a structural requirements for neutrophil receptor interaction. Agents Actions Suppl 1991;35:17–21.
[44] Bubeck P, Grotzinger J, Winkler M, Kohl J, Wollmer A, Klos A, Bautsch W. Site-specific mutagenesis of residues in the human C5A anaphylatoxin which are involved in possible interaction with the C5A receptor. Eur J Biochem 1994;219:897–904.