

John D. Lambris<sup>○</sup>,  
 John Pappas<sup>○</sup>,  
 Manolis Mavroidis<sup>○</sup>,  
 Yang Wang<sup>○</sup>,  
 Holly Manzone<sup>○</sup>,  
 Joseph Swager<sup>▲</sup>,  
 Louis Du Pasquier<sup>▲</sup> and  
 Randi Silibovsky<sup>○</sup>

Departments of Pathology and  
 Laboratory Medicine, University of  
 Pennsylvania<sup>○</sup>, Philadelphia and  
 Basel Institute for Immunology<sup>▲</sup>,  
 Basel

## The third component of *Xenopus* complement: cDNA cloning, structural and functional analysis, and evidence for an alternate C3 transcript\*

Although the third component of complement has been purified from two amphibian species, *Xenopus laevis* and the axolotl, only limited information is available about its primary structure in these species. We now present (a) 95% of the cDNA sequence encoding C3 from a *Xenopus laevis/Xenopus gilli* (*Xenopus* LG) hybrid (b) an analysis of the C3 convertase and factor I cleavage sites in *Xenopus* C3, and (c) evidence for an alternative form of C3. The *Xenopus* LG sequence has a 57% nucleotide and 52% amino acid sequence identity to human C3 and contains one potential N-glycosylation site in the  $\beta$ -chain. The deduced amino acid sequence showed that the C3 convertase and factor I cleavage sites (Arg-Ser) are conserved in *Xenopus* C3 and protein sequencing of *Xenopus* C3 fragments fixed on zymosan during complement activation demonstrated that *Xenopus* C3 is indeed cleaved by C3 convertase and factor I at these sites. Our screening of a liver cDNA library identified an unusual C3 clone with a deletion of 2502 bp, suggesting the presence of a novel C3 transcript in *Xenopus* LG liver. The presence of this C3 transcript was confirmed by reverse transcription polymerase chain reaction using *Xenopus* LG liver mRNA and specific oligonucleotide probes. This transcript encoded a putative 102-kDa protein comprising the  $\beta$ -chain of C3, together with the first 59 residues and the last 103 residues of the  $\alpha$ -chain; it would therefore lack many of the ligand binding sites found in the intact  $\alpha$ -chain. However, the molecule may be an analog of a truncated C3 molecule that is found in the serum of allergic dermatitis patients and acts as an inhibitor of eosinophil cytotoxicity and neutrophil adherence.

### 1 Introduction

The third component of complement (C3) is an integral part of the complement cascade, participating in both the activation and regulation steps of this cascade [1, 2]. It is the most versatile of the complement proteins interacting with about 20 other components of the cascade [1, 2]. The C3 protein represents the meeting point in the classical and alternative pathways, where they converge to form the C5 convertase that leads to the initiation of membrane attack complex formation. The alternative pathway of complement is the phylogenetically older pathway and is present even in cyclostomes, which until now have not been proven to have antibodies [3]. In mammals the alternative pathway is composed of C3, factors B, D, I and H, and properdin. Although C3 has been purified from several representative classes of vertebrates [1, 4] only limited information exists

regarding the presence of other alternative pathway proteins. Functional assays have suggested the presence of factor I- and H-like proteins in various classes of vertebrates, including fish and amphibians [5, 6], and a cDNA sequence encoding a factor B-like protein in the lamprey has recently been obtained [7].

The amphibian complement system is similar to that of mammals in employing both the classical and alternative pathways [8, 9]. C3 purified from *Xenopus* [8, 9] and *Ambystoma* (axolotl) [10] plasma has an  $\alpha$ - $\beta$ -chain structure, contains a thioester bond in the  $\alpha$ -chain and is glycosylated in the  $\beta$ -chain. Receptors mediating phagocytosis of C3-coated particles have been described in *Xenopus* macrophages [11]. Compatibility studies with *Xenopus* C3 and human C3 binding proteins have shown that *Xenopus* C3 binds to human CR1 and CR2 but not to factor H [9], suggesting that the binding sites for some of these proteins are different in the two species. Previous attempts to obtain the complete cDNA sequence of amphibian C3 have been unsuccessful; only a 900-bp 3'-sequence of *X. laevis* has been reported [12]. In this study we have obtained 95% of the cDNA sequence encoding C3 from a *X. laevis/X. gilli* hybrid (*Xenopus* LG) and have compared the deduced primary structure of this gene product to those of C3 from other species and of other homologous proteins. We have also characterized the factor I and C3 convertase cleavage sites in *Xenopus* C3 and obtained evidence that *Xenopus* has proteins analogous to factors I, H, B and D. Finally, we have identified a novel C3 transcript encoding a truncated *Xenopus* C3, a molecule which may represent the analog of a truncated human C3 that inhibits eosinophil cytotoxicity and neutrophil adherence [13].

[I 13589]

\* This work was supported by National Science Foundation grants DCB90-18751 and MCB9319111, National Institutes of Health grant AI 30040 and Cancer and Diabetes Centers Core Support Grants CA 16520 and DK 19525.

**Correspondence:** John D. Lambris, Laboratory of Protein Chemistry, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104-6079, USA (Fax: 215/573-2059; E-mail: lambris @ MSCF.MED.UPENN.EDU)

**Key words:** Complement / *Xenopus* / C3 / cDNA cloning / Alternative transcripts

## 2 Materials and methods

### 2.1 Animals and reagents

Generation of the *X. laevis* and *X. gilli* (LG) hybrid has been described previously [12]. All LG animals were kept at the Basel Institute for Immunology. *Xenopus* livers were excised and either frozen in liquid nitrogen or used immediately to prepare the cDNA libraries. All chemicals for automated sequencing were obtained from Applied Biosystems (Foster City, CA). Restriction enzymes were purchased from Boehringer Mannheim (Indianapolis, IN) or Promega (Madison, WI) and all radionucleotides from NEN (Boston, MA). All other chemicals and reagents used were of reagent or higher grade.

### 2.2 Screening of the *Xenopus* LG cDNA library

An oligo-dT primed *Xenopus* LG liver cDNA library was constructed using an Amersham (Arlington, IL) cDNA synthesis kit according to the manufacturer's instructions. The initial screening of this library was performed using a 325-bp Tth111I fragment purified from the EcoRI insert of a *X. laevis* pC3-29 clone [12]. Subsequent screenings were performed using oligonucleotides designed on the basis of the 5'-sequence of the initial C3 clones. The probes were labeled either with [<sup>32</sup>P]dCTP by use of a random labeling kit (Boehringer Mannheim) or with [ $\gamma$ -<sup>32</sup>P]dATP by kinase end-labeling kit (Promega). Screening was performed in duplicate, and hybridization was carried out as previously described [14]. Initially, 80 000 plaques were screened using the 325 bp Tth111I fragment, and of the 20 clones analyzed, 3 (X9, X11 and X3) were found to be 0.4, 2.0, and 3.1 kb in length, respectively. The largest clone was completely sequenced, and an oligonucleotide from the 5'-end of this clone was synthesized for re-probing the library. A 1 kb C3 clone (XOS-1) was found to extend the sequence of X3 by about 800 bp. Additional screening with a 25-bp oligonucleotide probe from the 5'-end of the XOS-1 sequence led to the isolation of a 2.4 kb clone (Xu4), which was fully sequenced. The Xu4 sequence matched that of X3 and XOS-1, but extended 701 bp beyond the C3 sequence in the 5' direction and had a deletion of 2502 bp. The clones and sequencing strategy are shown in Fig. 1.

### 2.3 DNA sequencing

Positive C3 clones were subjected to *in vivo* excision by helper phage R408 and recircularized to generate subclones in the Bluescript SK (-) vector [15]. The inserts of

Bluescript clones were then sequenced using the universal M13 forward/reverse primers and synthesized oligonucleotides. All oligonucleotides were synthesized using a Milligene Cyclone oligonucleotide synthesizer (Millipore, Bedford, MA). Sequencing was performed according to the method of Sanger [16], using the Sequenase 2.0 kit from US Biochemical according to the manufacturer's instructions. Both cDNA strands were completely sequenced at least once. Sequences were compiled and analyzed using the ASSEMBLER program of PC Gene (Intelligenetics, Inc. Mountain View, CA). The protein sequences for human, rabbit, rat and mouse C3 were taken from Swiss Protein Sequence databases. The protein sequences for guinea pig, lamprey, cobra, trout and hagfish C3 were translated from their cDNA sequences taken from the Genbank Sequence databases. The chicken C3 sequence was obtained in this laboratory [17]. All other analyses were performed using PCGene software (Intelligenetics).

### 2.4 Isolation of *Xenopus* LG mRNA, reverse transcription polymerase chain reaction (RT-PCR), and Southern blotting

*Xenopus* LG liver total RNA was isolated by the guanidinium isothiocyanate method, followed by CsCl centrifugation [18]. Polyadenylated RNA was purified using the PolyA Tract mRNA isolation kit (Promega) according to the manufacturer's instructions. The first strand of cDNA synthesis was constructed using the Amersham cDNA synthesis kit and the oligonucleotide HM17 (5'-TGGTTCTCACGTTGTTG-3') (Fig. 5) as primer, according to the manufacturer's instructions. The RNA/cDNA pool was then subjected to PCR using the primers HM13, HM9 and MM2 (5'-CAATACAG-CTCAAAGATCGG-3', 5'-ATC-TCCACCTGCTCGTTC-3', 5'-CCCCAAATCAGATAA-TCTC-3', respectively). The PCR was carried out using 2.5 U Taq polymerase (Promega) as follows: 30 cycles of 95°C for 1 min, 50°C for 2 min, 72°C for 2 min and a final extension of 10 min at 72°C. The PCR products were electrophoresed in 1% agarose and transferred to nitrocellulose, and the blots were hybridized with kinase 5'-end-labeled oligonucleotide MM1 (5'-CTGGATGGAATGCAGGAGAACC-3'). The sequences of the labeled bands were obtained by isolating the individual bands and ligating them into a pCR II plasmid (Invitrogen, San Diego, CA).

A negative control sample containing all the reagents except the RNA was included in all experiments to check for contamination. As positive controls, PCR reactions were carried out with DNA from clone Xu4 as template.

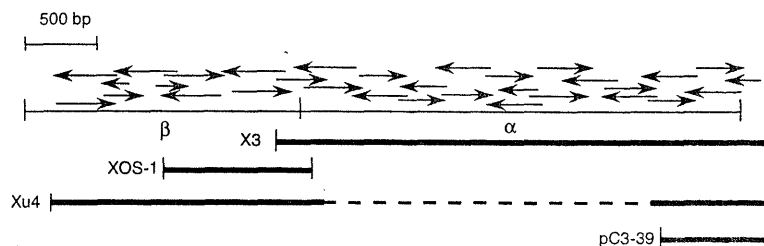


Figure 1. Sequencing strategy for C3 for *Xenopus* LG: Arrows show sequencing by oligonucleotides. cDNA clones that were isolated and sequenced are also shown.

CGCGCCGGTAAACATGGCTTCCTTGGGACGGCCACTGTAAAGATTCCTCTGGAGATT 60  
 A A G N N G F L G T A T V K I P S G D L 20  
 TCCEAAGATACAAAAGCAAGCAGTATGTATGTCTCAGCGTGAATCAATGTATGTCC 120  
 S K R D T K A S K Q Y V Y V S V K S N V C A 40  
 CTGGAGAAAGTGTGCTCTGTCTGACACAGTGGCTATATCTTCTTCCAGACTGACAA 180  
 L E K L V L L L S Y H S G I I S Q S H K L P 60  
 ACAATCTACACACCAGGATCTACAGTTTGTACCCGATTTACTCCATGAATACAGAGT 240  
 T I Y T P P G S T V L Y R I Y S M N Y K M 80  
 CAGCCTATTCCCAAGCCTCATCATTTGAATTTTACCCAGATGTTATTATCGTGAAC 300  
 Q P I S K T L E F L T P D G I T I V N 100  
 ACACATACCCTCCCAAGACGCAATCTGGTATCATCTCACAGTCCCACAAATCGCA 360  
 T H T I L Q D S K S G I I S Q S H K L P 120  
 GAACCTTTGAGCCTTGGAGTGTGGACTATCTCAGCAAAAGTATGAAGACACCCCTCAAC 420  
 E L L S L L G V W T I S A K Y E D T P Q Q 140  
 AATTACACCCTCACTTTGAGGTGAAGAAATATGTCTACCCCTCAATAGAAATTTCTG 480  
 N Y T I N F E V L K E Y V L P S I E I L 160  
 AAGCCAGACACTAATCTCTCTATAGTGTGCTGAATCTTCCGAGTGGATATTCAGGCC 540  
 K P D T N Y T S D A E S F G V D I Q A 180  
 CAGTATCTGTATGGGAACACGCTGGATGGATACCGCTTTGTGCTTTGGCATAATGAG 600  
 Q Y L Y G K H V D G Y A F V L F G I M K 200  
 GACAACTTAAGAAAGAAATCCCAAGTCCGCTGACAAAGATCAGGATTTGATGGAGAA 660  
 D N V K K G I P E S L T R I R I D D G E 220  
 GGTGCTGCTGAATCAAAAGGAAAGACCTAGTGAAGTCTTTGAGAAACAGAGGACATG 720  
 G R A E L K R K D L V K Y F E K P E E D M 240  
 CTACAATTCAGCTTGTATGTGACTGTCTGTGTGTGACAGAAACAGACTAGCAGCATGG 780  
 L Q F S L Y V T V S V V T E T G S D M V 260  
 GAGGCTGAGATGAGCATATCAGTATTTGTGACGACGCCATCAAAATCCCTTTACCAA 840  
 E A E I E H I S I V T T P Y K I L C T T K 280  
 ACCTCCAAATATTTCAAGCCTGGAAATCCCTTTGTATATGATGGTCTATGTAACCAATC 900  
 T S K Y F K P G M P F D M M V Y V T N P 300  
 GACCGTCTCCTGCTCTGCTATCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960  
 D G S P A R R I R V V V E P G N V E G S 320  
 ACTCAAGGAGATGGAACCAAGGCTGACACTGAACACACCCGCAATATAAACCTTTG 1020  
 T Q G G T R L T L N T P A N I N L L 340  
 CCCATACTGTAAACCAAGGACCCCACTTTGCTCCCGGCGCCCAAGCAATGCCACC 1080  
 P I T V K T R D P T L P P G R Q A T A T 360  
 ATGACTGCTACAGCTTACCAGCCCTCAAAGGACCAAGGAACTACCTGATATCAGCAAT 1140  
 M T A T A Y R P S K A Q G N Y L H I S I 380  
 GCTGGTCCAGATATAAACCTGGAGAAACATCCCTGTGAACCTTAAATTCGGAACCACT 1200  
 A G S D I K P N F N E N I P V N I R N T 400  
 GATGCGGCGTCCAGAACCAATAGAACATTTACATATCTGATCATAAGCAGGGGAAG 1260  
 D A G G R Q K Q I E H F T Y L I I S R G R 420  
 ATTCTAAGGTGGGAGGCGACGACCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1320  
 I L K V G R Q Q R Q P G Q P F V T M S I 440  
 CCTGTACCAGTACCTTGATACCATCTTTCCGATTTGTGGCATAATACATTGCTCAGT 1380  
 P V T D T I A T P R I V A Y Y I V S A 460  
 GGTGGAGGACGTGATATTTGATCTGACTCTCTTTGGGTGGATGGTGGTGGTGGTGGT 1440  
 G G G R D I V S D S L W V D V V D C M 480  
 GGCACATTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1500  
 G T L S V T G D K D R D N T I Q T P G S 500  
 CCAATGAGGCTGAAGCTGAGAGCAGACCACAGGCCTATGTGGGATTTGGTGGCTGGAT 1560  
 P M R L K L R L A D H K A Y V G L V A V D 520  
 AAGGTTCTATGCTGCTGACAGTAAATTAAGATCACACAAAGAAAGTATGGGATCA 1620  
 K G V Y V L N S K F K I T Q K R V W D S 540  
 GTGGAGAGTCCAGCATTTGGTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1680  
 V E K S D I G C T P G S G A N S E G V F 560  
 TCTGACGCTGGCTTGTCTACAGACGCTCTTTGGTATCAATACAGTCAAAGATCCGGAT 1740  
 S D A G L A L Q I N T A Q R S D 580  
 GCTCAATCTCCAGCCAGCAGCAAGCAGAAAACCGTCTCATCTGTTGCACTCATTGAGAT 1800  
 A Q C P A P A Q R R K R S S V A L I E I 600  
 AAATCCCGCAAGCTTCAGATACAGGACAGGCAAGGCAAGTCTGCTGGATGGAAATG 1860  
 K S G K A S E Y K D K A K K C C L D G M 620  
 CAGGAGAACTGATGGGCCACACCTGTGACCGCAGCGCCGCTACATCTTGGATGGGAAA 1920  
 Q E N L M G H T C D R R A R Y I L D G K X 640  
 GAATGTGGATGCTTTTGTGACTGCTGCAATATCTATGAGAAGAAGAGGAAGCTGAG 1980  
 E C V D A F V D C C K Y Y E K R E A E 660  
 AAGATTAGTAAGATGATGACCCCTTGGAAAGATGATGAAGATGACGAATACATGCTG 2040  
 K I S K D D D T L G R S D E D E Y M L 680  
 GATTCGACATTTGCTCCAGGACAGATTTCTGAGAGCTGGTCTGGAAGGTGGAAACA 2100  
 D S D I V S R T E F P E S W F W K V E Q 700  
 ATGGTTGAAAAGCCTGATGCAACGGTATTTCAAGCAAGCCCTAAATGTTTTCTTAAAA 2160  
 M V E K P D V N G I S S K T L N V F L K 720  
 TATTCATCACCTGGGAAAGTTCTGGCTGTGAGTTTGTCTGAAAACAAAGGCTCTGT 2220  
 D S I T C T W E V L A V S L S E N K G L C 740  
 GTGGTCCAGCCTATGAGATTAAGTGTGAGGATTTCTTATCGACCTGAAACTTCCA 2280  
 V G Q P Y E I K V M K D F I D L K A L P 760  
 TATTCTGTTGTGAGGAACGAGCAGGTGGAGATTCGCGCATTCTCTACAATACAGGAAT 2340  
 Y S V V R N E Q V E I R A I L Y N Y R N 780  
 GACAGGATTAAGGTGCTGGAGCTGATCCACAATCCAGAGTTTTCAGTCTGTCCAG 2400  
 D R I K V R V E L I H N P E F C S L S 800  
 GCTAAGAAGATACAGACAGGAGTTGGATGTGAGCCCTCTCTCCATCGCTGCCA 2460  
 A K K K Y R Q G E V W I G A L S I A V P 820  
 ATTGTCATTGTGCCACTCCTTGGCCAGCATGACATAGAGGTGAAGCATCTGTTTCA 2520  
 I V I V P L T L G Q H D I E V K A S V S 840  
 GCTCAATCTGGATCTTTGGTGAGATGGTGTCCGTAAAAGCTGAAAGTGTGGCTGAA 2580  
 A Q C S G F F G A D G V R K K L K V V P E 860  
 GGAATGGCATTCAAGATGTAAGACTGTATATTAGACGGCGAAGTTAAAGGGAAA 2640  
 G M R I A Q D V K T V I L E C P E V K G K 880  
 GATGGCGTCAAGAAGAAAAGTGAAGCACTGAATCCAAAAATGTTGTCGGCGAACC 2700  
 D G V Q E E K V E A L N P K N V V P R T 900  
 GACATTTGACACCACCTACTTTGCAAGAAACCCGATCAGTCAAGTGGTGAAGATGCC 2760  
 D I D T T I T L Q G T P I S Q M V E D A 920  
 ATTGATGGCAAAATATGATCATCTGATTTGGTTCGGCTGGCTGGGGAGCAGAA 2820  
 I D G N N M N H L I V V P A G C G E Q N 940  
 ATGATCTCAACGACTCTAGTGTCAITGGCCACGCTACCTGGATGCAAGTGGCCAAAT 2880  
 M I S T T P S V I A T R Y L D A S G Q W 960  
 GAGCGTGTGGAGTGAACCCGAGACCAAGCCCTCAAAAACATGAGACAAAGCTATGCT 2940  
 E R V G N R R D Q A L K W L I L E K Q K P 980  
 CAACAAATGGCTTTCCGCAACAGCAACTCTCAGCAGCTGGAAGGATGAGCCAGCG 3000  
 Q Q M A F R K P D N S Y A G C K D R P A 1000  
 AGTACGTGGCTCACAGCTATGGCTAAAGTGTGGGATGGCTCAGGAATTTATTGAT 3060  
 S T W L T G Y V A K V F G M A Q E F I D 1020  
 ATTGAGCTAATGTTCTCTGCGGCTCCGTCAAATGGCTTACTGAGGAAACAGAAAGCA 3120  
 I E A N V L C G S V K W L I L E K Q K P 1040  
 GATGGATTTGTCAGGAGATGCGCCATCCATCAGGAGATGGTGGGAGGATCACA 3180  
 D G L G T Q E N A P V I H Q E M V V G G I T 1060  
 ACAGGACGACGAGGGAGACTCTCCCTCAGCCCTTTATCTTATCCGCTATGTAGAA 3240  
 T G A A E G D S S L T A F I L I A M L E 1080  
 TGCCAAAGACTTGTATGAGCATGTGAACAATCTCAAGTCAGTATTGATAAGCCCTCA 3300  
 C Q R T C N E H V N N L Q V S T K N S 1100  
 AGTTACCTGGTGGTCAATATCCAGGACTAAAAAACCAATATCCATGCAATCACTTCC 3360  
 S Y L L V G Q Y P G L K K P Y A S I A I T S 1120  
 TATGCTTTCAGCTGCGGAAAGCTTCCAAACACAAATAGTGTCTGCTCATCTATA 3420  
 Y A L A L A G K L P N T N K L L S A S I 1140  
 GGTAAACCCACTGGGAGGACCGGAAACGCTTTATCTCCCTGGAACCATCATAT 3480  
 G N T H W E E P G K W L I L E C T S Y 1160  
 GCCCTCTAACCTCTTAAAAATGAAGAAATTTGATCTAACAGGTGGCATTGTACGTGG 3540  
 A L L T L L K M K E F D L T G G I V R W 1180  
 CTCAAATGAGCAGATACCTATGGCGAGTTTATCCACCCAGGCAACCATCTGATG 3600  
 L N E Q R Y Y G A V Y G S T Q A T I V M 1200  
 TTCCAAGCTTGTCTCAGTACCAGACAGATATCCCAAGTGTCAATGAGTGAACCTTAG 3660  
 F Q A L A Q Y Q T D I P S N A E L N L D 1220  
 GTCTCCCTCATCTCCAGAGGCAACAACCTTAACTATCTGATCATCATGAAAT 3720  
 S L H L P E R Q Q P L T Y R I N H E N 1240  
 GCTTTGCTGGCAGCTTCTGAGAGACTAGTTAAACCAAGACTTTGGTAAAGCTAAA 3780  
 A L L A R S A E T R L N O D F V V K A K 1260  
 GGGAAAGGCAAGGTACAATGAGGCTGGTACCTCTACCATGCTTGTGACAGGAAA 3840  
 G K G Q G T M R V V T V Y H A L V E K 1280  
 GAAGAAGCTTCACTTCACTGTGTGAAGTGGAGAGTAGAATTCGCAAGA 3900  
 E R K C L N F H L S V K V E V E L A R 1300  
 CCTCCGATGGTCAAAGGCAACAGTGTCTATAGAAGCTGTCGAAGGACTCAAGAAG 3960  
 P P D G A K A T V S I E A C A R H L K N 1320  
 GTGATGCCACCATTCATATTTGATATTTCCATGATGACTGGCTTTCTCTGATACC 4020  
 V D A T M S I D I S M T T G F S P D T 1340  
 AATTCCCTGGATAGCTAATGAAGGAGTGGCAAAATACATCTCAATATGAAAGTCAAC 4080  
 N S L D R L M K G V D K Y I S K Y E V N 1360  
 AAAGAGCAAAATGACAAGGCAACTTCTCTACTTGGAAAGTCTCAGACATGAG 4140  
 K G A N D K G T L I L Y L D K V S H I E 1380  
 GAAGAATGTGGAAGTTTATGCTCATCAATCTTGAAGTGGFTTTCATCCAGCCAGCT 4200  
 E E C V F Y A H Q Y T F V G G F I K 1400  
 TCTGTAAGTGTATGACTATTATATCCAGATAATCGCTGCACTAAATTTTACCATGT 4260  
 S V T V Y D Y Y T P D N R C T K F Y H V 1420  
 GAAGAAGGCAAGTGGTGGGAGGATTTGCAAGGAGTATGCGCGTGTGCAAGAA 4320  
 E E G S A L L G R I C Q G D L C R C A E 1440  
 GAGAATTTTATGACGAGCAAAATGAGGTTAAAATCACTGCTGAAATGAGAATCAAC 4380  
 E N C F M Q Q I E G K I T A E M R I N 1460  
 ATGGGTTGCGCTCCCGAGTGGATTTGTGTAAGGCTACTCTCAGGAAAGTGCAGCC 4440  
 M G C A P G V D F V Y K A T L T E V Q P 1480  
 AGTGACAATTTGACAATCTGTTATGACATTAAGAAGTCTAAGAAGGCAAGAT 4500  
 S D N F D N Y V M T I K K V I K Q G T D 1500  
 GAGGACTCTGAGGACAGCAGTAAATTTATCAGCCATATCAATCCCGAAAGCTTTA 4560  
 E D P E D K T R N F I S H I K C R K A L 1520  
 AATATGACGCTGAACCCAGATATCTGATTTGGGGGTTAAGTGGTGGTGGTGGTGGT 4620  
 N M Q L N R D Y L I W G V T G D L W R Q 1540  
 CCAGATGGATTTCTCAATCATTTGGGAGGACACATGGATGGAGTGGTGGCCATGAG 4680  
 P D G Y S Y I I G K D T W M E W W P N E 1560  
 AGGGATGTCACACGCTGAGAACCCAGGATCTCTGCGATGACTTTGAGACAGTTTCT 4740  
 R E C Q R E N Q D L C D D F E T V S D 1580  
 AACCTGGAGATTTGTTGTTTCCCAACTSAGGAACCTAAATTCAGACCAAAATGTTTC 4800  
 N L E I V G C P N -  
 CATCTTCCCAAGCTCATCTGTGATGTTTAAATTTGTTGTTGTTGTTGTTGTTGTTG 4880  
 GAAATTCGATATCAGCTTATCGATACGTCGACTCCGACCC 4921

Figure 2. Nucleotide and amino acid sequence of *Xenopus* LG C3: Double-underlined sequences show sequences determined by Edman degradation of zymosan-eluted C3 fragments. Single underlining shows the sequence of truncated C3. The Arg-Arg-Lys-Arg linker sequence is indicated by\* and the putative splice site by x.

## 2.5 Amino acid sequence determination of the C3 convertase and factor I-generated C3 fragments

*Xenopus* C3 fragments generated by *Xenopus* C3 convertase and factor I were prepared by incubating serum with zymosan as previously described [14]: *Xenopus laevis* serum (5 mg/ml) was incubated for 30 min at 25°C, with zymosan particles, which had been boiled in 2% SDS and then extensively washed with 2% SDS. The  $\beta$ -chain and the equivalent to the 43-kDa C-terminal fragment of the human C3 were eluted from the zymosan by boiling for 5 min in electrophoresis reducing sample buffer, and the  $\alpha'$ -chain and/or its 68-kDa N-terminal fragment were then eluted with 0.1 M NaHCO<sub>3</sub> containing 0.1 M hydrazine, (pH 11.0) for 60 min at 37°C [19]. As a control we used human serum treated in the same manner as the *Xenopus* serum, except for being incubated at 37°C instead of 25°C. The supernatants from both incubations were analyzed by 10% SDS-PAGE [20], and the individual bands were sequenced on an Applied Biosystem 473A gas-phase protein sequencer as previously described [21].

## 3 Results

### 3.1 Isolation and nucleotide sequencing of cDNA encoding C3 of *Xenopus* LG

We have previously reported [12] the partial sequence (analogous to residues 1321-1641 of human C3 sequence) of *X. laevis*. All previous attempts to obtain the full *Xenopus* C3 sequence were unsuccessful because the libraries we have generated from *Xenopus laevis* contained only short inserts [12]. In order to obtain additional *Xenopus* C3 sequence we first prepared a new liver cDNA library using the vector  $\lambda$ -ZAP II and screened this library with a 325-bp Tth111I fragment purified from the EcoRI insert of pC3-39 clone [12]. We then analyzed 20 out of 130 positive clones obtained and subsequently generated nested deletions and obtained the full sequence of the largest clone, the 3.1 kb clone X3. Using the 5'-terminal sequence of this clone as a basis, we prepared an oligonucleotide probe that was used to rescreen the library and the remaining clones from the earlier screening with the Tth111I fragment. We found one clone (XOS-1) that extended the sequence of X3 by about 800 bp. Additional screening with a 25-base oligonucleotide probe synthesized from the 5'-terminal sequence of XOS-1 led to the isolation of a 2.4 kb clone (Xu4), that extended the sequence of XOS-1 by about 701 bp (see below for more information on

Xu4). The compiled coding sequence was found to be 4770 bp in length (Fig. 2). The deduced amino acid sequence of the *Xenopus* C3 was 31-54% identical that of C3 from other species (Table 1).

The deduced amino acid sequence indicated an  $\alpha$ -chain of 997 amino acids (112.5 kDa). The highly conserved Arg-Arg-Arg-Arg linker sequence that links the  $\beta$ -chain to the  $\alpha$ -chain in proC3 of most species [14] was found to Arg-Arg-Lys-Arg in *Xenopus* LG. A single glycosylation site was identified at position 141 (N-Y-T).

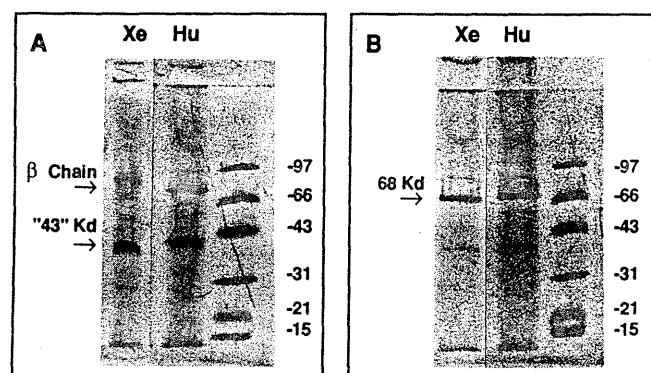
### 3.2 The C3 convertase and factor I cleavage sites in *Xenopus* C3

To confirm that *Xenopus* C3 is indeed cleaved at the C3 convertase and factor I cleavage sites predicted from the deduced cDNA sequence, we generated C3 fragments (C3b/iC3b) by activating *Xenopus* serum with zymosan, analyzed the fragments by SDS-PAGE and obtained their N-terminal amino acid sequences.

Electrophoretic analysis of the C3 fragments fixed on zymosan after activation of *Xenopus* serum showed a degradation pattern similar to that observed for human C3 (Fig. 3). Whereas the same fragments were generated in the presence and the absence of EGTA, no fragments were detected when the activation was performed in the presence of EDTA. These observations suggest that *Xenopus* has proteins with functions similar to those of human factors B, D, I and H. The N-terminal amino acid sequence of the 68-kDa fragment (Fig. 2) showed that this fragment is the analog of the human C3 fragment generated by C3 convertase and the cleavage site for the C3 convertase is conserved in *Xenopus* C3. Furthermore, the N-terminal sequence of the 43-kDa fragment (Fig. 2) indicated that this fragment is the analog of the human C3 fragment generated by factors I and H (Fig. 3) and that the factor I cleavage site is Arg-Ser, as in the C3 molecules of most other species [14].

**Table 1.** Sequence similarity of *Xenopus* C3 to C3 from other species

Species	% identity	% similarity
Cobra	54.2	67.7
Mouse	52.6	67.1
Guinea pig	52.4	66.6
Rat	52.0	66.7
Human	50.9	65.3
Trout	45.1	59.4
Lamprey	33.2	49.9
Hagfish	31.1	48.5



**Figure 3.** SDS-PAGE electrophoresis of *Xenopus laevis* and human fragments eluted from zymosan. The zymosan-bound C3 fragments were prepared as described in Sect. 2.5. (A) Fragments eluted after reduction of zymosan-iC3b with 2-ME (B). Fragments eluted after treatment of zymosan-iC3b with hydrazine (first preeluted as in A).

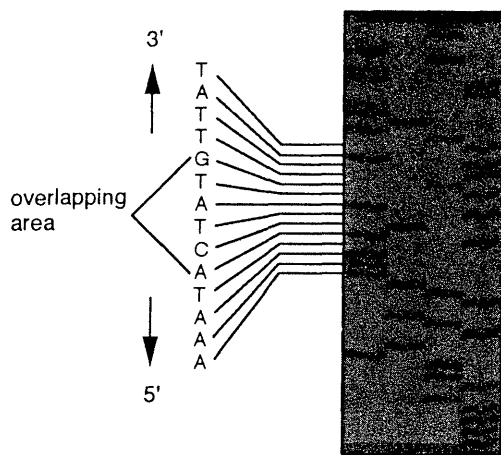


Figure 4. Nucleotide sequence of the splice site of truncated C3. The sequencing gel and the corresponding nucleic acid sequence of the positive strand are shown. Both the positive and negative strands were sequenced.

### 3.3 Characterization of a truncated transcript of *Xenopus* LG C3

When we screened the cDNA library with an oligonucleotide probe derived from the 5'-terminal sequence of the XOS-1 clone, we isolated a 2.4 kb clone (Xu4). The sequence of Xu4 matched that of X3 and XOS-1, but extended 701 bp beyond the C3 sequence in 5'-direction (Figs. 1 and 2) and had a deletion of 2502 bases. This deletion, which did not alter the reading frame, corresponded to amino acid residues 731-1560 (according to human C3 numbering). The exact splice position could not be determined because there was a six-nucleotide overlap between the donor and acceptor sites (Fig. 4). The deduced sequence of the truncated C3 included the entire  $\beta$  chain, 59 amino acids extending from the N-terminus of the  $\alpha$ -chain (*i.e.* C3a minus 18 amino acids from the C-terminus) and 103 amino acids of the C-terminus of the  $\alpha$ -chain and the Arg-Arg-Lys-Arg processing site. This finding suggests that the CR1 binding domain (which

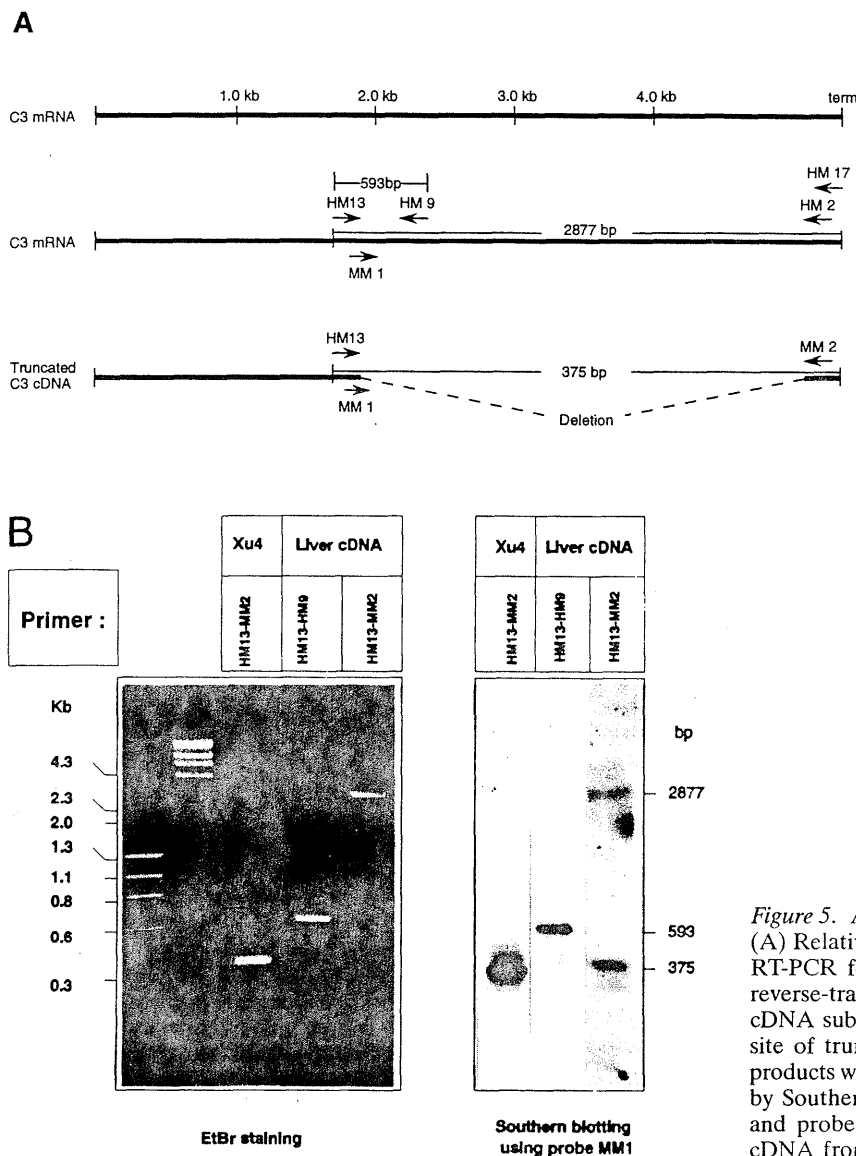


Figure 5. Analysis of C3 transcripts in *Xenopus* LG liver. (A) Relative position of primers and of C3 transcripts. (B) RT-PCR for C3 transcripts. Samples of liver RNA were reverse-transcribed using primer HM17 and the resultant cDNA subjected to PCR using primers flanking the splice site of truncated C3 as described in Sect. 2.4. The PCR products were electrophoresed in agarose gels and analyzed by Southern blotting with probe MM1 as indicated. PCR and probe hybridization controls were established using cDNA from clone Xu4.

includes the CR2, factor B and factor H sites) at residues 727-768 [22], the CR2 binding site at residues 1199-1210 [23], factor H binding domain at 1187-1249 [24], and the properdin binding site at 1402-1435 [25] were all deleted. Assuming that no other alterations had occurred in the 5'-end of the unsequenced segment, the expected size of this alternative form of C3 was about 102 kDa.

To ascertain the presence of truncated C3 in the original library and to exclude any *in vivo* excision artifacts, DNA was isolated from the  $\lambda$ -ZAP cDNA library and subjected to PCR using the oligonucleotides HM13 and MM2 from areas adjacent to the deletion segment (Fig. 5). A fragment of the expected size was found, indicating that the library did contain the deletion clone (data not shown). To confirm the presence of truncated C3 in the liver, poly(A<sup>+</sup>) RNA was subjected to reverse transcription and amplification by PCR using specific oligonucleotide primers flanking (HM13 and MM2), and within (HM9), the deleted segment (Fig. 5). The PCR products were resolved by agarose gel electrophoresis and analyzed by Southern blotting with a radiolabeled oligonucleotide probe that specifically hybridizes with both C3 and truncated C3 (MM1). The expected transcripts of 2877 bp and 375 bp were obtained with primers HM13 and MM2, and one of 593 bp with primers HM13 and HM9 (Fig. 5). These data indicate that both normal and truncated C3 are encoded by *Xenopus* LG liver. The specificity of the RT-PCR was demonstrated when DNA from clone Xu4 was used as templates and hybridized with probe MM1 (Fig. 5). Primers HM13 and MM2 generated the expected 375-bp signal with Xu4. The nature of the 375-bp product produced using primers HM13 and MM2 and liver RNA was further analyzed by isolating this band, subcloning it into the pCRII plasmid and sequencing it. The sequence obtained exactly matched the nucleotide sequence of the deletion clone.

#### 4 Discussion

In our investigations we have extended the characterization of the primary structure of *Xenopus* C3. In addition, we have presented evidence for the presence of proteins analogous to factors I, H, B and D in *Xenopus* and have characterized a novel C3 transcript in this species. The compiled sequence that we have obtained represents 95% of the *Xenopus* C3 sequence. Repeated attempts to complete the sequence of *Xenopus* C3 by screening our library were unsuccessful, most probably because of the length of the C3 mRNA and the fact that the library we used was generated by oligo-dT priming. The obtained sequence, however, contains all the information necessary to characterize the known functional sites of C3 [14]. The identity of the deduced amino acid sequence to those of C3 molecules from other species ranged from 31% to 54%. Whereas the linker that connects the  $\alpha$ - and  $\beta$ -chains of proC3 is composed of four arginine residues in most species [14], the linker in *Xenopus* C3 was Arg-Arg-Lys-Arg, as in hagfish C3 [9]. Despite this substitution, the overall charge was preserved, suggesting that this arrangement is important for proper processing of proC3 into a mature two-chain molecule. We found only one potential glycosylation site in the  $\beta$ -chain, which had previously been shown by [<sup>125</sup>I]Con A binding studies to be glycosylated [9]. As

predicted, the  $\alpha$ -chain had no glycosylation sites, and its estimated molecular mass of 112.5 kDa was in agree with 112 kDa previously obtained by SDS-PAGE [26].

The deduced amino acid sequence of residues 1299-1641 (according to human C3 numbering) from the previously published cDNA sequence for *X. laevis* C3 (4490) between residues 1299-1641 apparently contains three additional cysteines when compared to the corresponding segment of *Xenopus* LG. We have now repeated the sequencing of *X. laevis* clone C3-39 containing the additional cysteines and found that the sequencing gel contains two compressions. Sequencing using ITP clearly resolved the sequence of this segment and showed its complete identity with that of *Xenopus* LG, thus demonstrating that there are no extra cysteines in this segment of *Xenopus* C3. All other cysteines are completely conserved in the C3 of *X. laevis* and the LG hybrid.

The C3 convertase cleavage site (residues 726-727 in human C3), which is an Arg-Ser in all known species except lamprey [27], was conserved in the *Xenopus* LG sequence. The cleavage site that corresponds to the second factor I cleavage site of human C3 (residues 1298-1299) was also conserved (Arg-Ser), in contrast to that corresponding to the first factor I cleavage site (residues 1281-1282), which was Arg-Glu instead of Arg-Ser. N-terminal sequencing of C3 fragments eluted from zymosan confirmed that *Xenopus* C3 is indeed cleaved at the sites predicted to correspond to the C3 convertase and second factor I cleavage sites. Thus far, we have no evidence that the *Xenopus* C3 is cleaved at the Arg-Glu bond; further kinetic experiments are in progress to clarify this issue. Generation of fragments similar to those produced by mammalian factors I, H and C3 convertase suggests that *Xenopus* has proteins with similar activities. The recent cloning of a cDNA sequence encoding a protein with high sequence similarity to mammalian factor B has confirmed the presence of a factor B analog in *Xenopus* [28].

We have also isolated a cDNA clone with a sequence identical to that encoding the full *Xenopus* C3, except for a deletion of 2502 bases from the  $\alpha$ -chain; the truncated cDNA encoded a putative protein of about 102 kDa. The existence of the truncated cDNA was confirmed by RT-PCR, and Southern blot analysis and sequencing established that the fragments we obtained were of the lengths expected (375 bp and 2877 bp for the truncated and intact C3 using primers HM13 and MM2 and 593 bp for intact C3 with primers HM13 and HM9).

The existence of a form of C3 comprised of a  $\beta$ -chain and small segments of the  $\alpha$ -chain is particularly interesting in light of the findings of Minkoff and co-workers [13]. These researchers have isolated a protein that acts as an inhibitor of eosinophil cytotoxicity and neutrophil adherence from the serum of patients with allergic dermatitis. Not only was the N-terminus of their protein identical to that of the  $\beta$ -chain of human C3, but the  $\beta$ -chain of human C3 had the same activity as the isolated serum factor. The mechanism by which this  $\beta$ -chain analog is produced is unclear. Because the isolated  $\beta$ -chain is insoluble and the purified factor had a molecular weight higher than that of the human  $\beta$ -chain, we hypothesize that it may contain portions of the  $\alpha$ -chain attached to the  $\beta$ -chain. Generation of an alternate

form of C3 such as we have described for *Xenopus* could provide a mechanism by which such a protein is produced.

Until now, only one other alternative transcript for C3 has been described [29]. This transcript encodes a product that is thought to be involved in B cell activation. However, this transcript differs from the one described in the present study in that it encodes only the C-terminal portion of the  $\alpha$ -chain, whereas the alternative transcript we describe encodes the  $\beta$ -chain and small segments of the  $\alpha$ -chain. In light of their possible immunologic applications, the structure and function of these C3-related molecules merit continued investigation.

We thank Dr. D. McClellan for editorial assistance and Mrs. Jian Pang for her excellent technical assistance. The nucleotide sequence data reported in this paper have been submitted to the GenBank Nucleotide Sequence Database under the accession number U19253.

Received September 29, 1994; accepted December 8, 1994.

## 5 References

- 1 Lambris, J. D., *Immunol. Today* 1988. 9: 387.
- 2 Müller-Eberhard, H. J., *Annu. Rev. Biochem.* 1988. 57: 321.
- 3 Ishiguro, H., Kobayashi, K., Suzuki, M., Titani, K., Tomonaga, S. and Kurosawa, Y., *EMBO J.* 1992. 11: 829.
- 4 Lambris, J. D., *Complement Profiles* 1993. 1: 16.
- 5 Kaidoh, T. and Gigli, I., *J. Immunol.* 1989. 142: 1605.
- 6 Kunnathmuglia, L. M., Chang, G. H., Sim, R. B., Day, A. J. and Ezekowitz, R. A. B., *Mol. Immunol.* 1993. 30: 1249.
- 7 Nonaka, M., Takahashi, M. and Sasaki, M., *J. Immunol.* 1994. 152: 2263.
- 8 Sekizawa, A., Fujii, T. and Katagiri, C., *J. Immunol.* 1994. 133: 1436.
- 9 Alsenz, J., Avila, D., Huemer, H., Esparza, I., Becherer, D., Kinoshita, T., Wang, Y., Oppermann, S. and Lambris, J. D., *Dev. Comp. Immunol.* 1992. 16: 63.
- 10 Avila, D. and Lambris, J. D., *Comp. Biochem. Physiol.* 1990. 95B: 839.
- 11 Sekizawa, A., Fujii, T. and Tochinai, S., *J. Immunol.* 1984. 133: 1431.
- 12 Grossberger, D., Marcuz, A., Du Pasquier, L. and Lambris, J. D., *Proc. Natl. Acad. Sci. USA* 1989. 86: 1323.
- 13 Minkoff, M. S., Wong, W. W. and Silberstein, D. S., *J. Exp. Med.* 1991. 174: 1267.
- 14 Lambris, J. D., Lao, Z., Pang, J. and Alsenz, J., *J. Immunol.* 1993. 151: 6123.
- 15 Short, M. G., Fernandez, M. J., Sorge, A. J. and Huse, D. W., *Nucleic Acids Res.* 1988. 16: 7583.
- 16 Sanger, F. S., Nicklen, S. and Coulson, A. R., *Proc. Natl. Acad. Sci. USA* 1977. 74: 5463.
- 17 Mavroidis, M., Sunyer, J. O. and Lambris, J. D., *J. Immunol.* (in press) 1995.
- 18 Maniatis, T., Fritsch, E. F. and Sambrook, J., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989.
- 19 Law, S. K. and Levine, R. P., *Proc. Natl. Acad. Sci. USA* 1977. 74: 2701.
- 20 Laemmli, U. K., *Nature* 1970. 227: 680.
- 21 Becherer, J. D. and Lambris, J. D., *J. Biol. Chem.* 1988. 263: 14586.
- 22 Becherer, J. D., Alsenz, J., Esparza, I., Hack, C. E. and Lambris, J. D., *Biochemistry* 1992. 31: 1787.
- 23 Lambris, J. D., Ganu, V. S., Hirani, S. and Müller-Eberhard, H. J., *Proc. Natl. Acad. Sci. USA* 1985. 82: 4235.
- 24 Lambris, J. D., Avila, D., Becherer, J. D. and Müller-Eberhard, H. J., *J. Biol. Chem.* 1988. 263: 12147.
- 25 Daoudaki, M. E., Becherer, J. D. and Lambris, J. D., *J. Immunol.* 1988. 140: 1577.
- 26 Corbi, A. L., Miller, L. J., O'Connor, K., Larson, R. S. and Springer, T. A., *EMBO J.* 1987. 6: 4023.
- 27 Nonaka, M. and Takahashi, M., *J. Immunol.* 1992. 148: 3290.
- 28 Kato, Y., Saltercid, L., Frajnik, M. F., Kasahara, M., Namikawa, C., Sasaki, M. and Nonaka, M., *J. Immunol.* 1994. 153: 4546.
- 29 Lernhardt, W., Raschke, W. C. and Melchers, F., *Curr. Top. Microbiol. Immunol.* 1986. 132: 98.