The Rat and Mouse Homologues of MASP-2 and MAp19, components of the Lectin Activation Pathway of Complement

Cordula M. Stover, Steffen Thiel, Nicholas J. Lynch and Wilhelm J. Schwaeble

*J Immunol* 1999; 163:6848-6859; 
http://www.jimmunol.org/content/163/12/6848
The Rat and Mouse Homologues of MASP-2 and MAp19, components of the Lectin Activation Pathway of Complement1,2,3

Cordula M. Stover,* Steffen Thiel, † Nicholas J. Lynch,‡ and Wilhelm J. Schwaeble4,*†

Recently, we described two novel constituents of the multimolecular initiation complex of the mannan-binding lectin (MBL) pathway of complement activation, a serine protease of 76 kDa, termed MASP-2, and a MASP-2 related plasma protein of 19 kDa, termed MAp19. Upon activation of the MBL/MAp19 complex, MASP-2 cleaves the fourth complement component C4, while the role of MAp19 within the MBL/MAp19/MASP-2/MAp19 complex remains to be clarified. In humans, the mRNA species encoding MASP-2 (2.6 kb) and MAp19 (1.0 kb) arise by an alternative polyadenylation/splicing mechanism from a single structural MASP-2 gene. Here, we report the complete primary structures of the rat homologue of MASP-2 and of rat and mouse MAp19. We show that both MASP-2 and MAp19 are part of the rat MBL pathway activation complex and demonstrate their exclusively hepatic biosynthesis. Southern blot and PCR analyses of rat genomic DNA indicate that as in humans, rat MASP-2 and MAp19 are encoded by a single structural gene. *The Journal of Immunology, 1999, 163: 6848–6859.

Within the host’s innate immune defense, the mannan-binding lectin (MBL)-mediated activation of complement may be especially important to deal with encapsulated microorganisms that are able to evade Ab recognition and alternative pathway activation (1). By virtue of its carbohydrate recognition domains, MBL is able to bind to carbohydrates on the surfaces of pathogens (2) and opsonize these for uptake by receptors (3, 4); via the association with its serine proteases MASP-1 (93 kDa) (5) and MASP-2 (76 kDa) (6), MBL is able to activate complement (7). Formation of the C3 convertase is initiated by MASP-2-mediated cleavage of C4 (6), while the C2 cleaving enzyme of the lectin activation complex has not yet been defined.

1 Abbreviations used in this paper: MBL, mannan-binding lectin; MASP, MBL-associated serine protease; MAp19, MBL-associated plasma protein of 19 kDa; CUB, C1r/C1s/Uegf/bone morphogenetic protein 1; EGF, Epidermal growth factor; CCP, complement control protein; ORF, open reading frame; UT region, untranslated region.

Materials and Methods

Materials

Restriction enzymes were purchased from Boehringer Mannheim (Mannheim, Germany). Two different rat liver cDNA libraries constructed from normal and acute-phase liver RNA (catalog nos. 937506 and 936512) and a mouse liver cDNA library constructed from liver RNA of a B6CBAF1/J mouse (catalog no. 935302) were purchased from Stratagene (Cambridge, U.K.). Poly(A)Tract mRNA Isolation System IV was obtained from Promega (Madison, WI). RT-PCR kit (Superscript preamplification system)
was obtained from Life Technologies (Paisley, U.K.). [α-32P]dCTP, blotting membranes, and protein weight markers were from Amersham-Pharmacia (Uppsala, Sweden). PCR2.1 vector and prokaryotic expression vector pTrxFus were purchased from Invitrogen (Leek, The Netherlands).

**Isolation of MASP-2 and MASP19-specific cDNA transcripts from rat liver cDNA libraries**

A partial cDNA clone specific for rat MASP-2 (RT-PCR rl-1, EMBL Nucleotide Sequence Database accession no. Y18285) was generated by RT-PCR from rat liver RNA using randomized primers as previously described (8). After EcoRI excision from PCRII (TA cloning kit; Invitrogen), the 623-bp fragment was [32P]dCTP labeled by random priming (Random Primed Labeling Kit; Boehringer Mannheim) and used as a probe to screen two rat liver cDNA libraries cloned in the phage vector lambda ZAP (Stratagene). Of 2 × 10^7 plaques of each cDNA library, 18 hybridizing phages were obtained and subjected to another round of screening. Plasmid pBluescript SK+ was rescued by in vivo excision. Sequence analysis revealed that both clones were partial transcripts of the mouse MASP-2 mRNA with clone pml-1 (12) representing the complete 5′ sequence of this mRNA species. Therefore, pml-1 was sequenced in full.

**Cloning of partial cDNA transcripts of rat MASP-2 by reverse transcriptional amplification**

Rat liver RNA was reverse transcribed using Superscript premolution system (Life Technologies). The oligo(T)-primed cDNA obtained was used in cDNA amplifications as follows. Oligonucleotides were derived from the 5′ untranslated (UT) region of clone pml-1, a complete transcript of the 1.0-kb MAp19 mRNA species (5′-GGG ACA CAT ATG AGG CTA-3′, position 4–21), and from sequence coding for the C-terminal portion of the CUB domain of clone pml-1, one of the partial transcripts of the 2.6-kb MASP-2 mRNA species (5′-GGC GGA TGT TGT AGG CCG-3′, position 447–464). A 658-bp amplification product was obtained. RT-PCR, rtM2, subcloned in PCR2.1 (TA cloning kit; Invitrogen) and sequenced on both strands.

**Analysis of the genomic organization of rat MAp19**

Genomic DNA was extracted from tissues (spleen, testis, submandibular glands, and lungs) of an adult Wistar rat according to the standard Proteinase K preparation method (11). Approximately 15 µg were digested overnight with EcoRI or PstI, separated on an 0.8% (w/v) agarose gel, blotted to a nitrocellulose membrane (BDH, Poole, U.K.), and hybridized according to standard protocols (13), and mRNA was purified using Poly(A)Tract mRNA Isolation System IV (Promega). Approximately 2 µg of Poly(A+) RNA per lane was separated on a denaturing 6% (w/v) agarose gel and transferred to Hybond N membrane (Amersham-Pharmacia). The blots were hybridized according to standard protocols (11) with [35S]riboprobe-labeled rat MASP-2 cDNA probe (pml-1, EcoRI/XhoI excised). To validate Northern blot results, the membranes were rehybridized with a [32P]dCTP-labeled β-actin cDNA probe (14).

Total RNA was also extracted from cultivated splenocyte-derived cell lines (rat hematopoietic cell line JBI and rat liver epithelial cell line RLE). The cells were stimulated with various cytokines (TNF-α, IL-1β, and IFN-γ (Sigma-Aldrich, St. Louis, MO), and IL-6 (R&D Systems, Minneapolis, MN); 10 ng/ml each (final concentration); 24 h). Total RNA was separated on a denaturing agarose gel, blotted, and hybridized with radio-labeled cDNA clone pml-1 and washes as previously described (11). The membranes were exposed to x-ray film (blue sensitive: Genetic Research Instrumentation, Dunmow, U.K.) for 14 days at −70°C using intensifying screens.

**Characterization of the 1.0-kb MAp19 mRNA of the mouse**

The rat MAp19 cDNA sequence showed high score of identity with a partial murine cDNA sequence contained in the mouse expression sequence tags gene bank library (GenBank). This clone (I.M.A.G.E. Consortium clone ID 1297088) (12) was supplied to us by the HGMP Resource Centre (Cambridge, U.K.) and further characterized by Southern blotting (using clone pml-5 as a probe) and sequence analysis.

**Isolation of MASP-2-specific cDNA transcripts from a mouse liver cDNA library**

Using radiolabeled pml-1 as a probe, we screened ∼2 × 10^7 plaques of a mouse liver cDNA library cloned in the phage vector XZap (Stratagene) by cross-hybridization. After rescreen, two potential mouse MASP-2-specific cDNA clones were obtained of which the plasmid pBluescript SK+ was rescued by in vivo excision. Sequence analysis revealed that both clones were partial transcripts of the mouse MASP-2 mRNA with clone pml-1 (12) representing the complete 5′ sequence of this mRNA species. Therefore, pml-1 was sequenced in full.

**Northern blot analysis**

Total RNA was isolated from various tissues from healthy adult Wistar rats (brains, hearts, livers, lungs, kidneys, spleens) according to standard protocols (13), and mRNA was purified using Poly(A)Tract mRNA Isolation System IV (Promega). Approximately 2 µg of Poly(A+) RNA per lane was separated on a denaturing 6% (w/v) agarose gel and transferred to Hybond N membrane (Amersham-Pharmacia). The blots were hybridized according to standard protocols (11) with [35S]riboprobe-labeled rat MASP-2 cDNA probe (pml-1, EcoRI/XhoI excised). To validate Northern blot results, the membranes were rehybridized with a [32P]dCTP-labeled β-actin cDNA probe (14).

**In situ hybridization**

Sections (14 µm thick) were cut from snap-frozen rat embryos (gestational age, 16 days), fixed in phosphate-buffered formaldehyde, and prehybridized as described previously (15). 32P-labeled cRNA was transcribed in ammonium and sense orientation from a subfragment of rat MASP-2 cDNA cloned in pBluescript KS+ representing the serine protease domain and the 3′ UT region specific for the 2.6-kb MASP-2 mRNA species (XhoI/EcoI0109 I, clone pml-1, 486 bp) and from a subfragment of rat MAp19 cDNA cloned in pBluescript SK+ representing the CUB I domain (pml-5 religated after removal of 3′ EcoI0109 I fragment, 338 bp). After application of specific [35S]riboprobes, incubation at 65°C for 16 h, and stringent washing (15), radioactive signals were detected by autoradiography on Hyperfilm β-max (Amersham-Pharmacia).

**Prokaryotic expression of recombinant rat MAp19**

After cyclic amplification of clone pml-5 using BamHI-modified sense oligonucleotide 5′-GGG ATC CCA CAC TTT TGG CCT CCA AG-3′ (position 593–612) and Universal primer (5′-ACT GCA AGA ATC TTG GTA GCT GCA CCT CCA AG-3′, position 354–368) synthesized on both strands, clones were cut expression vector pTrxFus (Invitrogen). Electrocompetent GI1724 Escherichia coli host was transformed, and ampicillin-resistant colonies were analyzed by colony blotting and hybridization with [32P]dCTP-labeled pml-1 cDNA. Several recombinants were assayed for protein production according to the manufacturer’s protocol. As a control, noncyclic amplified DNA was transferred to the coding sequence with the expression vector pTrxFus alone which was included in the experiment. In brief, recombinants were grown to OD 600 nm 0.5 at 30°C, then tryptophan (Sigma-Aldrich; 0.1 mg/ml final concentration) was added and cultures were shaken at 37°C for 4 h (200 rpm). Cells were harvested by centrifugation (2000 × g, 30 min, 4°C) and recombinant protein released by sonication in the presence of 0.5% SDS (Soniprep 150; Sanyo, Gallenkamp plc, U.K.). Protein expression was analyzed by 15% (v/v) SDS-PAGE (reducing conditions) and Western blotting using rabbit anti-human MASP-2 antiserum produced by immunizing rabbits with recombinant MASP-2 expressed in E. coli (8).
Detection of MASP-2 and MAp19 as constituents of the MBL pathway initiation complex in rat plasma

Mannose was coupled to TSK-75 beads (Merck, Darmstadt, Germany) as previously described (16). Then, 500 μl Mannose-TSK beads were mixed with 4 ml rat plasma and 26.5 ml TBS/Tween/Ca²⁺ (10 mM Tris, 150 mM NaCl, 0.1% NaN₃, pH 7.4, 0.5% (v/v) Tween 20, 5 mM CaCl₂) and incubated overnight at 4°C. The beads were washed with TBS/Tween/NaCl and then 500 μl TBS with 20 mM EDTA and 375 μl SDS-PAGE sample buffer were added to the beads. Half of the material was reduced by adding DTT to 0.6 M followed by boiling for 3 min, and then iodoacetamide was added to 0.7 M. The other half iodoacetamide only was added to 0.7 M followed by boiling for 3 min. The samples were next loaded in 5-cm wide wells, and the proteins were separated by SDS-PAGE on 4–20% gradient gels followed by blotting onto polyvinylidene difluoride membranes (Hybond-P, Amersham-Pharmacia Biotech). After blotting, the membranes were blocked by incubation in TBS with 0.1% (v/v) Tween 20, 1 mg/ml human serum albumin, and 100 μg/ml human IgG, washed, and developed by enhanced chemiluminescence detection (Pierce, Rockford, IL).

Results

Characterization of rat MASP-2 and rat MAp19 mRNA

To determine the nucleotide sequences of the two mRNA species that strongly hybridized with a partial cDNA clone (clone RT-PCR rl-1) specific for rat MASP-2 in a Northern blot analysis of rat liver RNA (8), two different rat liver cDNA libraries (both from Fisher 344 rats, one constructed of normal, one of acute-phase liver tissue RNA) were screened by filter hybridization using clone RT-PCR rl-1. In total, nine potential cDNA clones for rat MASP-2 and/or MAp19 mRNA were sequenced in full. As shown in Fig. 1A, five partial cDNA transcripts (ranging from 1.0 to 1.9 kb in length) of the 2.6-kb rat MASP-2 mRNA were obtained. The longest cDNA transcript, clone prl-8 (1397 bp) has an open reading frame (ORF) of 1928 bp, a polyadenylation initiation signal (AATAAA), followed by a translation termination codon (TTA) and a short 3’ UT region. Clone prl-8 overlaps with the four shorter transcripts obtained for rat MASP-2 mRNA, clones prl-1 (1898 bp), prl-3 (1657 bp), prl-9 (1112 bp), and prl-2 (1027 bp). The overlapping sequences of these clones revealed complete identity. The uninterrupted ORF at the 5’ end and the comparison to the human cDNA sequence for MASP-2 showed that none of these clones represented the complete MASP-2 mRNA sequence. In the human, the 5’ ends of MASP-2 and MAp19 mRNAs are identical (8). Therefore, to obtain the missing 5’ portion of rat MASP-2 mRNA, an RT-PCR reaction was performed on rat liver RNA using an oligonucleotide derived from the known 5’ sequence of rat MAp19 mRNA (see below) and an oligonucleotide corresponding to the coding sequence for CUB domain II (represented in MASP-2, not in MAp19), as described in Materials and Methods. A cDNA transcript of 658 bp, rM2 (RT-5’), was amplified (see Fig. 1A). The sequence of rM2 (RT-5’) overlaps with 522 bp of the 5’ end of prl-8 and extends it by 136 bp. It comprises the complete 5’ coding sequence of the rat MASP-2 2.6-kb mRNA preceded by a 5’ UT sequence of 9 bp. The composite sequence of the rat MASP-2 mRNA is shown in Fig. 1B.

The derived amino acid sequences of rat MASP-2 and human MASP-2 (8) share an overall identity of 79.9% (Fig. 2). Cysteine residues that are important for the secondary structure of the proteins (17–19) are conserved in all structural domains, as well as the arginine and isoleucine residues at the predicted cleavage site to separate the A chain from the B chain upon activation, and also the two juxtaposed dyads, serine and histidine, histidine and aspartic acid, respectively, shown to be important in the catalytic function of serine proteases (20) (Fig. 2). An octamer motif found between the CCP domain II and the serine protease domain of human MASP-2, PVRCEPVCV (amino acid position 413–420), (21) is conserved in the rat (PVRCKPVVC) and indicated in Fig. 2. By analogy to human MASP-2 and MAp19 (where the N terminus of the mature proteins was determined by peptide sequencing (6, 8)), the mRNA species for rat MASP-2 and rat MAp19 encode a putative leader peptide of 15 aa (see Fig. 1B and Fig. 3B). The start of the respective mature rat plasma proteins MASP-2 and MAp19 is predicted to be TLLGSKWPEP. The calculated molecular mass of mature rat MASP-2 is 73954 Da.

Four overlapping clones specific for rat MAp19 mRNA were sequenced, clones prl-5 (716 bp), prl-6 (251 bp), prl-7 (680 bp), and prl-4 (969 bp) (Fig. 3A). Clones prl-6 and prl-7 showed to be partial transcripts of the 1.0 kb MAp19 mRNA, while clone prl-5 proved to be a full-length cDNA transcript comprising 12 bp of the 5’ UT region, followed by the complete coding sequence (558 bp), including a translation termination codon (TAA) and a 3’ UT region of 144 bp (including a polyadenylation initiation signal (AATAAA) and a Poly(A) tail) (Fig. 3B). The overlapping sequences of all transcripts obtained for rat MAp19 mRNA were identical. The 5’ UT sequence of prl-5 is identical with the 5’ UT sequence of the MASP-2 mRNA; however, the 3’ UT regions of MASP-2 and MAp19 mRNA are different. Like human MAp19, rat MAp19 comprises two structural motifs, an N-terminal CUB domain and an EGF-like domain followed by a C-terminal sequence of 4 aa (not contained in the deduced amino acid sequence for MASP-2) (glutamic acid, glutamine, serine, and leucine). The calculated molecular mass of mature rat MAp19 is 19318 Da.

Characterization of the genomic organization of rat MAp19

Another cDNA transcript isolated, prl-4, differs from the MAp19-specific cDNA clones prl-5, -6, and -7 by having a 444-bp insertion between the coding sequence for the EGF-like domain and the coding sequence for the unique carboxyl terminus of MAp19 (EQSL) (Fig. 3A). To assess the relative abundance of this transcript in rat liver RNA, a RT-PCR was performed using oligonucleotides flanking the inserted sequence of 444 bp (see Materials and Methods). The 222-bp product, rM2 (RT-PCR) contained only sequence specific for clone prl-5 (Fig. 3A). No signal was detected on Northern blot analysis of rat liver RNA, using a probe specific for this inserted sequence (data not shown).

To establish whether the 444-bp insertion in prl-4 is due to a splicing intermediary that retains intronic sequence (22), a PCR analysis of rat genomic DNA was performed using MAp19-specific oligonucleotides flanking this presumed intronic sequence as described in Materials and Methods. A 620-bp product (pgr-5) was obtained and subcloned in PCR 2.1. Sequence analysis of this clone revealed complete identity between the genomic amplification product and the corresponding region of clone prl-4 (data not shown), suggesting the view that prl-4 retained an intronic sequence. Sequence alignment between the intron separating the coding sequence for the EGF-like domain and the corresponding intron in rodent MAp19 (see Materials and Methods) (Fig. 4, A–C). The genomic sequence obtained comprises 1798 nt (pgMAp19) and was in full agreement
with nucleotide sequences obtained from overlapping parts in cDNA clones prl-1, prl-3, prl-2, prl-8, prl-9, RT-5' rM2, and RT-PCR rl-1) of the rat homologue of the 2.6-kb MASP-2 mRNA species. Indicated restriction sites were used to establish subclones for sequence analysis. Arrows indicate the reading of sequencing reactions performed. B. The primary structure of rat MASP-2 mRNA composed of overlapping sequences of cDNA clones prl-1, prl-3, prl-2, prl-8, prl-9, RT-5' rM2, and RT-PCR rl-1. All clones were sequenced in full. Sequence comparison revealed 100% identity in areas of overlap. The derived amino acid sequence is given beneath. The first position of the mature MASP-2 protein is indicated (+1). Nucleotide and amino acid positions are numbered. Polyadenylation initiation signal (AATAAA) within the 3' coding sequence is underlined.
(exon b). Approximately 800 bp downstream, one exon encodes the EGF-like domain (exon c), which is followed by the intron sequence of 444 bp characterized above (prl-4 and pgr-5). Exon d contains the coding sequence for the tetrapeptide sequence specific for the C terminus of MAp19 and sequence for the 3' UTR region of the 1.0-kb MAp19 mRNA species. There are no EcoRI or PstI sites contained in this genomic sequence. Splice junctions and intronic splicing enhancers, such as branchpoint sequences (23) and polypyrimidine rich tracts, 5' of the acceptor site, are indicated (Fig. 4D). There are two potential branchpoint sequences in intron c/d. The coding sequence for glutamic acid, glutamine, serine, and leucine 3' of the acceptor site matches a consensus sequence for an exonic splicing enhancer motif (24). As illustrated in Fig. 4E, classical splice sites (following the "gtag rule") allow joining of the exons a, b, c, and d. A strong polyadenylation initiation signal (AATAAA) is available for 3' modification of the primary transcript.

Southern blot analysis of rat genomic DNA using cDNA probes specific for MASP-2, MAp19, and the intronic sequence contained in clones prl-4 and pgr-5

The genomic organization of rat MASP-2 and rat MAp19 was further assessed by Southern blot analysis of restriction-digested rat genomic DNA. There are EcoRI or PstI sites contained in this genomic sequence. Splice junctions and intronic splicing enhancers, such as branchpoint sequences (23) and polypyrimidine rich tracts, 5' of the acceptor site, are indicated (Fig. 4D). There are two potential branchpoint sequences in intron c/d. The coding sequence for glutamic acid, glutamine, serine, and leucine 3' of the acceptor site matches a consensus sequence for an exonic splicing enhancer motif (24). As illustrated in Fig. 4E, classical splice sites (following the "gtag rule") allow joining of the exons a, b, c, and d. A strong polyadenylation initiation signal (AATAAA) is available for 3' modification of the primary transcript.

Primary structure of the mouse 1.0-kb MAp19 mRNA

Sequence analysis of I.M.A.G.E. Consortium cDNA clone 1297088 (Fig. 6A) revealed the complete primary structure of the mouse homologue of the 1.0-kb MAp19 mRNA. As shown in Fig. 6B, I.M.A.G.E. Consortium cDNA clone 1297088 represents 74 bp with the 5' end of the MAp19 mRNA hybridizes with a single fragment of 2.8 kb only (Fig. 5D). The same band hybridizes when a duplicate filter was probed with a radiolabeled fragment specific for the intron sequence located between coding sequences for the EGF-like domain of MASP-2/MAp19 and for the MAp19-specific C terminus (Fig. 5E).

One possible interpretation of these results is that rat MASP-2 and rat MAp19 are encoded by a stretch of genomic DNA that is 15 kb in length, although we cannot exclude the presence of a large intron that is flanked (and therefore excised) by restriction sites used. The predicted cleavage site used upon activation into the arginine and isoleucine residues is indicated by an arrow. An octamer motif observed between CCP domain II and the catalytic domain of the serine proteases C1r, C1s, MASP-1, and MASP-2 (21) is conserved in rat MASP-2 (underlined consensus). The translational products for human and rat MASP-2 show an overall identity of 79.9%.

FIGURE 2. Alignment of the deduced amino acid sequences of human and rat MASP-2 (hMASP-2/rMASP-2). Amino acids are numbered as of the mature start of the human MASP-2 protein, asterisks mark the signal peptide (6), and conserved residues, as found also in C1r, C1s, and MASP-1, are indicated: cysteines (bold) and the three residues essential for the catalytic sites of these serine proteases, histidine, aspartic acid, and serine (L). The predicted cleavage site used upon activation into the arginine and isoleucine residues is indicated by an arrow. An octamer motif observed between CCP domain II and the catalytic domain of the serine proteases C1r, C1s, MASP-1, and MASP-2 (21) is conserved in rat MASP-2 (underlined consensus). The translational products for human and rat MASP-2 show an overall identity of 79.9%.
the ORF encoding the first seven residues of the signal peptide (deduced in analogy to human MASP-2/Map19) differed from a recently published cDNA sequence for mouse MASP-2 mRNA (28). As we have shown in human and in rat that both MASP-2 and MAP19 mRNA species share an identical 5′ sequence, we aimed to clarify whether this divergence is a consistent feature in the mouse. Therefore, we isolated another mouse MASP-2-specific cDNA transcript, termed pml-1, from a mouse liver cDNA library.
Sequence analysis of clone pml-1 (1145 bp) revealed that it represents a 5' UT sequence followed by the coding sequence (ORF 1101 bp) for the putative signal peptide, the CUB I domain, the EGF-like motif, the CUB II domain, and the CCP I motif of mouse MASP-2. Fig. 6C shows a sequence comparison between I.M.A.G.E. Consortium cDNA clone 1297088 (encoding mouse MAP19), the corresponding 5' sequences of the mouse MASP-2-specific cDNA transcripts pml-1, and the previously published mouse MASP-2 cDNA sequence (accession no. AB009459, Ref. 28). It can be seen that both our mouse MAP19 and our MASP-2 specific cDNA transcripts share an identical 5' sequence that diverges from the corresponding sequence in accession no.

**FIGURE 4.** Organization and primary structure of the genomic DNA-encoding rat MAP19. A, Schematic presentation of the split gene organization of rat MAP19. B, Genomic fragments were amplified by PCR from rat genomic DNA using MAP19 cDNA and intron-specific oligonucleotides. The location of amplification products is schematically presented; the numbers indicate the nucleotide positions in cDNA clone prl-5 to which the primers anneal. C, Sequencing strategy. D, Primary structure of the genomic DNA coding for the rat plasma protein MAP19. Numbers indicate bp position. Intron sequences are in lower case, gt and ag, in-frame dinucleotide consensus splice sites; nucleotide sequences conforming to pre-mRNA consensus splicing signals (5', 3' splice sites, branchpoint sequences, polypyrimidine tracts) are indicated (25); 5' and 3' splice sites and polypyrimidine tracts are underlined, branchpoint sequences are in italic and bold. Note that there are two sequences in intron c/d that match the mammalian branchpoint consensus (26, 27). In exon d, the conserved coding sequence for the unique C terminus of MAP19 complies with a consensus sequence obtained for exonic splicing enhancer motifs (AGCAGAGCCTCT, marked with dots) (24). AATAAA, polyadenylation initiation signal; A, Poly(A) addition site of the rat MAP19 mRNA. E, Schematic presentation of posttranscriptional splicing of the rat MAP19 gene.
AB009459. As shown in Fig. 6C, the divergence includes the 5' UT sequence as well as the 5' coding sequence for the N-terminal part of the signal peptide with one variant (I.M.A.G.E. Consortium cDNA clone 1297088 and clone pml-1) encoding a signal peptide of 20 aa and the other variant (AB009459) encoding a signal peptide of 15 aa residues only. As all of our clones were isolated from the strain C57BL6 × CBA and clone AB009459 was derived from the BALB/C strain, we assessed whether the divergent 5' sequences were due to polymorphic variants among these mouse strains. An RT-PCR analysis was conducted on liver RNA preparations of both strains. One set of primers used was specific for the 5' UT sequence of I.M.A.G.E. Consortium clone 1297088 and clone pml-1 (5'-CCC ATA GTG CCT GGG G-3', position 1–16 of clone pml-1) combined with primers in antisense orientation specific for either mouse MASP-2 (5' -CAC CTG GCC TGA ACA AAG G-3', position 605–623 of clone pml-1) or for the 3' UT region of mouse MAp19 (5'-CGG AGG CAG GTT TAT TAT TG-3', position 3.4 kb only). D, Radiolabeled MAp19 cDNA clone pml-5 hybridizes with a 2.8-kb band. E, A radiolabeled fragment specific for the intron sequence located between coding sequences for the EGF-like domain of MASP-2/MAp19 and for the MAp19-specific C terminus (intron c/d, see Fig. 4) hybridizes with a 2.8-kb band.

Characterization of the site of biosynthesis for MASP-2 and MAp19 mRNA

MASP-2 and MAp19 mRNA expression was analyzed in Northern blot analysis of Poly(A+)-selected mRNA extracted from brains, hearts, livers, spleens, and kidneys of male and female rats. As shown in Fig. 7A, strong hybridization signals were seen in both liver RNA preparations, while in this and a series of other experiments (using different RNA preparations of the same organs and lung tissue in addition), no extrahepatic expression of MASP-2 or MAp19 mRNA could be detected. However, in liver both MASP-2 mRNA (2.6 kb) and an abundant MAp19 mRNA (1.0 kb) were consistently seen. Using MASP-2 and MAp19 mRNA-specific primers (5'-GGC TGA TGG ATT CTG GAC GAG-3', bp position 1052–1072 and 5'-TGC AAA TTA GAA ATT ATT TAT TAT-3', bp position 1857–1880 of clone pml-1 and 5'-TTC ACA GGA TTT GAG GCC TTC-3', bp position 391–411 and 5'-GTT GTCTGA CTC ACT GGG GT-3', bp position 593–612 of clone pml-5, respectively), a 800-bp and a 220-bp amplification product, respectively, were obtained by RT-PCR from rat liver RNA, while all other RNA preparations tested (brain, kidney, spleen, thymus) were negative (data not shown). The exclusively hepatic expression of MASP-2 and MAp19 mRNA is underlined by in situ hybridization results on whole-body cryostat sections of rat embryos. A riboprobe transcribed in antisense orientation from cDNA coding for the CUB I domain (Fig. 7C) and a riboprobe transcribed in antisense orientation from cDNA coding for the serine protease domain (data not shown) strongly hybridize with hepatic tissue only. On consecutive sections, no other organ showed any detectable level of MASP-2/MAp19 mRNA expression. The same observation was made on whole-body sections of rat embryos at an earlier stage of gestation (day 6), where only the hepatic Anlage strongly hybridized (data not shown). On Northern blot analysis of rat embryonic and postnatal liver RNA preparations (gestational age 15 days, and 1, 2, 14, and 18 days postnatally) showed consistent MASP-2 and MAp19 mRNA expression (data not shown).
FIGURE 6. Primary structure of the mouse MAp19 mRNA species and alignment of the 5' region of mouse MAp19 mRNA with the corresponding 5' region of mouse MASP-2-specific cDNA clone pml-1 and the previously published divergent 5' region for mouse MASP-2 generated by rapid amplification of cDNA end (28). Alignment of the deduced amino acid sequence of human, rat, and mouse MAp19 (hMAp19/rMAp19/mMAp19).

A, Schematic presentation of I.M.A.G.E. Consortium clone ID 1297088 representing mouse MAp19 mRNA and of pml-1, a partial mouse MASP-2 mRNA transcript and sequencing strategy. PstI and SacI were used for restriction mapping.

B, Nucleotide sequence of mouse MAp19 mRNA as represented by the I.M.A.G.E. Consortium clone ID 1297088. The derived amino acid sequence is given beneath. The nucleotide sequence and the derived amino acid sequence are numbered. In analogy to human MASP-2, the presumed signal peptide sequence is indicated with asterisks. The polyadenylation signal (aataaa) is in italic and underlined.

C, The nucleotide sequences of the 5' ends of mouse MAp19 cDNA (I.M.A.G.E. Consortium clone ID 1297088) and mouse MASP-2 cDNA (clone pml-1) are aligned, and their shared 5' UT region as well as identical sequence coding for the signal peptide are compared with those of the recently published mouse MASP-2 cDNA (28). Derived amino acids are given for each variant. The two signal peptides are marked with asterisks, their divergent N termini are in italic, and the start of the mature protein is numbered.

D, Amino acids are numbered as of the mature start of the human MAp19 protein, domain boundaries are indicated. The C-terminal sequence (EQSL) described for human MAp19 is conserved in rat and mouse MAp19. The translational products for human and rat MAp19 show an overall identity of 77.6%, those for human and mouse show an overall identity of 81.2%, and those for rat and mouse MAp19 show an overall identity of 94.1%.
However, in two hepatic cell lines (the rat liver epithelial cell line RLE and the rat hepatoma cell line JBI), no constitutive or cytokine-inducible (IFN-γ, TNF-α, IL-6, and IL-1β were tested) MASP-2/MAp19 mRNA expression could be detected (data not shown).

Characterization of MASP-2 and MAp19 in rat plasma

To determine whether anti-human MASP-2 antiserum cross-reacted with rat MAP19, we expressed recombinant rat MAP19 in the Thiofusion Expression System (Invitrogen) and analyzed the fusion protein in a Western blot analysis. As shown in Fig. 8A, the fusion protein composed of the mature sequence for rat MAP19 and the thioredoxin fusion partner stained strongly with our polyclonal anti-human MASP-2 rabbit antiserum (Fig. 8A, lane 2), while the fusion partner alone (Fig. 8A, lane 1) showed only weak staining. Thus, this Ab could be used to determine the presence of MASP-2 and MAP19 in rat plasma.

A preparation of MBL and MASPs was made by affinity purification of rat plasma on mannose-derivatized beads. When this preparation was analyzed in its nonreduced form, the Ab stained two bands representing the disulfide linked A and B chains of MASP-2 (76 kDa) and MAP19 (19 kDa) (Fig. 8B, lane 2). When analyzed after reduction, an additional band is seen (Fig. 8B, lane 4). In analogy to the human MASP-2 (6, 8), the upper 80-kDa band most likely represents unactivated MASP-2 in which no cleavage of the polypeptide chain has occurred. The middle band (60 kDa) represents the A chain of cleaved MASP-2, whereas the lower band represents MAP19. A similar pattern of bands is seen when human plasma is analyzed in this way (8).

**Discussion**

This contribution describes the primary structure, biosynthesis, and in part the genomic organization of MASP-2 and MAP19 in the rat. As in the human, MASP-2 and MAP19 are encoded by two mRNA species of ~2.6 kb and ~1.0 kb, respectively.
MASP-2 and MAp19 mRNA of the rat share an identical 5' UT region followed by an identical stretch of coding sequence comprising 543 bp. In addition, we also characterized a cDNA transcript specific for mouse MAp19. Surprisingly, the 5' sequence of mouse MAp19 mRNA differed considerably from the previously published cDNA sequence for mouse MASP-2 (28), which would imply that the organization of mouse MASP-2 and MAp19 differs from that of rat and human. However, isolation of another mouse MASP-2-specific cDNA transcript from a cDNA library revealed an identical 5' sequence to that obtained for our MAp19-specific mRNA transcript. This result was confirmed by RT-PCR analysis using sense primers derived from the 5' end of our MAp19- and MASP-2 specific cDNA sequences in combination with antisense oligonucleotides specific for either MAp19 or MASP-2 mRNA on liver RNA of two different mouse strains. To reveal the origin of the divergent 5' sequence, the RT-PCR analysis described above was repeated using a sense primer specific for this divergent 5' sequence in combination with either of the MASP-2 or MAp19 mRNA-specific antisense primers. Using this combination of primers, MASP-2 and MAp19 cDNAs were amplified, which demonstrates that in the mouse, two polymorphic forms of mRNA transcripts (leading to a polymorphism in the signal peptide) are present for MASP-2 mRNA as well as for MAp19 mRNA. The origin of this polymorphism observed for both MASP-2 mRNA and MAp19 mRNA in an inbred mouse strain remains to be clarified. There are at least two possible explanations: 1) a gene duplication event of the structural MASP-2 gene (from which both MASP-2 and MAp19 mRNA are generated by alternative splicing as in human and rat) has occurred in the mouse (leading to two closely related genes encoding either of the variant 5' sequences) or 2) a single structural mouse MASP-2 gene encodes two variants of both MASP-2 and MAp19 mRNA by an alternative splicing event in the 5' region in combination with the alternative splicing/polyadenylation event generating either MASP-2- or MAp19-encoding transcripts. In either case, the alternative usage of the polymorphic 5' sequences in both MASP-2- and MAp19-specific mRNA species supports the view that also in the mouse, MASP-2 and MAp19 mRNA are processed by alternative splicing/polyadenylation from a common primary gene transcript.

As a common feature, rat, mouse, and human MAp19 mRNA code for an identical C-terminal sequence (EQSL) followed by a stop codon and a 3' UT region (different from the 3' UT of MASP-2 mRNA) with a high degree of identity among these species. Recently, murine MASP-2 was mapped to the telomeric region of chromosome 4 (29), syntenic with human chromosome 1, on which we have located the MASP-2 gene at position 1p36.23-1p36.31 (30).

Based on the unexpectedly high degree of identity in the primary structure among species, we aimed to assess the genomic mechanism responsible for the generation of MASP-2 and MAp19 in the rat. Southern blot analysis of rat genomic DNA revealed that a probe representing a full-length transcript of rat MAp19 mRNA (prl-5) as well as a probe representing the intronic sequence retained in an incompletely spliced MAp19 cDNA clone (prol-4, intron c/d) strongly hybridize with a single PstI restriction fragment of 2.8 kb only. PCR amplification of rat genomic DNA revealed a partial gene structure of 1.8 kb comprising four exons encoding MAp19, which includes exon a (coding sequence for the signal peptide and the N-terminal portion of CUB I domain), exon b (coding sequence for the C-terminal portion of CUB I domain), exon c (coding sequence for the entire EGF-like domain), and exon d (coding sequence for the MAp19-specific C terminus and the 3' UT region of the MAp19 mRNA). The positions of the intron within the split exon for the CUB I domain and of the intron-separating coding sequence for the C-terminal CUB I domain and the EGF-like domain are conserved in human MASP-1 and human C1s, respectively (28). The intronic sequence located between sequence coding for the EGF-like domain and 3' sequence of MAp19 mRNA is identically positioned in human where it precedes the alternative splice/polyadenylation exon (exon b) to generate human MAp19 (8). The degree of identity between rat and human of this intronic sequence may suggest an important secondary structure necessary for spliceosome recognition (31).

In human and rat, nucleotide sequences presumably involved in the splicing/polyadenylation event are conserved. The striking identity of the nucleotide sequence coding for the unique C terminus of rat, human, and mouse MAp19 might underline its characteristics of a purine rich, exonic splicing enhancer shown to promote recognition of upstream splice signals by the spliceosome (24). Thus, the significance of this sequence as an RNA recognition element in the process of transcription may account for the cross-species conservation of the translational product derived from this coding sequence (32). Interestingly, this peptide adds strong hydrophilic charge to the C terminus of MAp19 as seen on Kyte-Doolittle analysis (not shown). The presence of an intron in conjunction with branchpoint, pyrimidine tract, and 3' splice site (criteria contained in both human and rat intron a/b, c/d, respectively) has been shown to function as an enhancer for recognition of a downstream poly(A) initiation signal (33). The poly(A) initiation signal for MAp19 mRNA is identical in all three species.

Thus, the Southern blot results in combination with the partial genomic analysis provide strong evidence that in the rat, as in human, both plasma proteins, MASP-2 and MAp19, originate from alternatively spliced/polyadenylated mRNA transcripts of a single structural gene.

Northern blot and in situ hybridization results indicate high tissue specificity for the biosynthesis of MASP-2 and MAp19 mRNA. In adult rats as well as in rat embryos, both mRNA species are exclusively expressed in hepatic tissue. This is in line with a recent finding showing that MASP-1 mRNA is also expressed in liver only and is not detectable in RNA preparations of human fetal heart, brain, lung, and kidney (34). Ito cells, the vitamin A-storing fat cells of the liver, were identified as the MASP-1-expressing cell type (35). MBL mRNA is expressed as early as 17 wk gestation in human fetal liver (36). In the adult, it may act as a moderate acute-phase reactant (36), increasing serum protein levels up to 3-fold (37). In humans and rodents, hepatic biosynthesis was shown for MBL (38-40). Recent reports suggest a pathogenic role of the lectin pathway of complement activation in IgA nephropathy (41, 42). Therefore, the liver may maintain a circulating pool of all components of the MBL-pathway initiation complex of complement. The hepatic biosynthesis may also contribute locally to the clearance of blood borne microorganisms (43, 44).

We have shown that both MASP-2 and MAp19 are constituents of the human and rat MBL pathway initiation complex of complement activation. Based on the homology between human, rat, and mouse MASP-2 and MAp19, we assume that both components are also constituents of the mouse MBL pathway initiation complex.

While the role of MASP-2 as the C4 cleaving enzyme of the MBL activation pathway is clear, the function of MAp19 in this complex remains to be clarified. A recent report by Takahashi et al.
(10) demonstrated a calcium-independent association between human MASP-1 and Map19. A subsequent report by Thiell et al. (Thiell, S., J. Petersen, T. Vorup-Jensen, M. Matsushita, T. Fujita, C. Stover, W. Schwaab, and J. Jensenius. 1995. Activation of complement by mannose-binding protein in association with a novel C1s-like serine protease. J. Exp. Med. 182:1563-1566) supported this hypothesis. In summary, we have demonstrated for the human and the rat that both plasma proteins MASP-2 and Map19 are generated by an alternative splicing/polyadenylation process from the same single structural gene and provided evidence that the same mechanism also accounts for both MASP-2 and Map19 in the mouse. The surprisingly high degree of similarity on the genomic, the mRNA, and the protein level between species as distant as human and rodents implies a selective pressure underlying such conservation (45).

Acknowledgments

We are most grateful to Drs. Keith Whaley, Ian Eperon, Kenneth B. M. Reid, and Robert B. Sim for their suggestions reading this draft. We thank Dr. Burkhard Schütz for his help in preparing rat embryos. We are grateful to Prof. Dr. Burkhard Schütz for his help in preparing rat embryos. We are grateful to Steen Vang Petersen. Ms. Kate Henderson is acknowledged for technical support. Computer based sequence alignments were kindly performed by Mr. Andrew Wallace. This manuscript is dedicated to Dr. Frank Harris, CBE.

References