Immunology 1998 **93** 421–430

Cloning and sequencing of a cDNA encoding chicken mannan-binding lectin (MBL) and comparison with mammalian analogues

S. B. LAURSEN,* T. S. DALGAARD,† S. THIEL, B.-L. LIM,‡ T. V. JENSEN,* H. R. JUUL-MADSEN,† A. TAKAHASHI,§ T. HAMANA,¶ M. KAWAKAMI¶ & J. C. JENSENIUS* **Department of Medical Microbiology and Immunology, University of Aarhus, Aarhus, Denmark, †National Institute of Animal Research, Foulum, Denmark, ‡Department of Zoology, University of Hong Kong, Hong Kong, §School of Fishery Science, Kitasato University, Sanriku, Japan, and ¶Fundamental Research Center, Kitasato Institute, Tokyo, Japan*

SUMMARY

The serum lectin, mannan-binding lectin (MBL) (also denoted mannan-binding protein or mannose-binding protein, MBP) has been identified in mammals (humans, monkey, cow, rabbit, mouse and rat). Upon binding to carbohydrates on the surface of microorganisms, MBL mediates activation of the complement system, leading to killing of the microorganism. MBL thus exerts a role in the innate immune defence. We have described the isolation and partial characterization of an analogous protein in chicken serum. Oligonucleotides based on the N-terminal sequence of this protein were used in a reverse transcription–polymerase chain reaction (RT–PCR) with chicken liver RNA as template. The PCR product was sequenced and found to encode part of the NH_2 terminus of chicken MBL. A perfect match probe was synthesized and used to screen a chicken liver cDNA library. The isolated clones carried a cDNA insert of 1692 bp with an open reading frame of 714 bp encoding a mature protein of 238 amino acids including a signal peptide of five amino acids. The deduced amino acid sequence agrees with those determined by conventional amino acid sequence analysis of the peptides except for four residues. We have compared the deduced primary structure of chicken MBL with the mammalian analogues. The phylogenetic analysis indicates that the gene duplication leading to two different MBL forms in mammals occurred after the split from birds and reptiles. This concurs with the finding of only one form of MBL in chickens.

INTRODUCTION

Mannan-binding lectin (MBL), also denoted mannose-binding protein or mannan-binding protein (MBP), has been identified in several mammalian species including humans,[1] monkey,[2] rabbit,[3] cow,[4] pig,[5] rat[6] and mouse.[7] An analogue protein has recently been found in chicken serum.[8,9]

MBL belongs to a family of proteins denoted collectins.[10] Other collectins are the serum proteins conglutinin and collectin-43 (CL-43) and the lung surfactant associated proteins SP-A and SP-D. The collectins belong to the superfamily of C-type lectins, characterized by the requirement of calcium ions for the binding of saccharides and a number of conserved amino acid residues in the globular carbohydrate recognition domain (CRD).[11,12]

Received 8 October 1997; revised 25 November 1997; accepted 3 December 1997.

*Department of Obstetrics and Gynaecology, Odense University Hospital, DK-5000 Odense C, Denmark.

Correspondence: Dr S. B. Laursen, Department of Obstetrics and Gynaecology, Odense University Hospital, DK-5000 Odense C, Denmark.

The collectins are oligomers of a characteristic subunit composed of three identical polypeptide chains. Each chain contains a short N-terminal region with two to three cysteine residues, followed by a collagenous region, containing the characteristic Gly-X-Y repeats with hydroxylated prolines and lysines, which allows three polypeptides to join in a collagen-like triple helical structure. A short 'neck region', involved in an α-helical coiled coil structure between three polypeptides,[13] is followed by the COOH-terminal CRD.

In monkey, mouse and rat two closely related MBLs have been identified called MBL-A and MBL-C whereas only one form has been found in humans.[1,2,7,14–16] From sequence data, it can be argued that human MBL most resembles the rodent C-form. Recently the cDNAs encoding both forms were described in monkey, whereas in apes (primates) only the MBL-C-like form could be shown.[2] Physiological data shows, that in the mouse only MBL-A mRNA but not the MBL-C mRNA is upregulated during inflammatory stimulation.[7] The concentration of MBL in human serum has been found to increase during acute-phase reactions[17] and this is supported by investigating the level of mRNA.[18,19] MBL activates the complement system, via the so-called MBLectin pathway,

independent of antibody but in a manner similar to that of the classic pathway. This activity has been demonstrated for human MBL and mouse MBL-A.[20–22] The fact that the human MBL amino acid sequence most resembles the MBL-C forms contradicts the functional data.

Analysis has shown a mosaic nature of the genes encoding human MBL (hMBL), rat MBL-A (rMBL-A) and rat MBL-C (rMBL-C), mouse MBL-A (mMBL-A) and MBL-C (mMBL-C), with each exon encoding a distinct domain indicating evolution of the gene through exon shuffling.[23–25] The hMBL and rMBL-A genes both contain four exons, with the first exon encoding the signal peptide, the N-terminal cysteine rich region and five to seven Gly-X-Y repeats. The rest of the Gly-X-Y repeats are encoded by exon two. The third exon encodes the 'neck region', a region forming an α-helical coiled coil conformation between three polypeptides and apparently involved in the initiation of the formation of and stabilization of the collagenous triple helices and finally, exon 4 encodes the entire COOH-terminal carbohydrate recognition domain (CRD) along with the 3′ UTR. The overall exon structure of mMBL-A (five exons), mMBL-C (six exons) and rMBL-C (six exons) resemble the hMBL-gene except for one (mMBL-A) or two (rMBL-C and mMBL-C) additional exons encoding parts of the 5′-UTR.[2,13,23–26]

The genes of the three known human collectins have all been assigned to the long arm of chromosome 10, where the MBL is closest to the centromere (10q21), followed by SP-D (10q22–23) and SP-A (10q22–23).[26,27] This close clustering of genes encoding similar proteins, suggests an evolutionary relationship. In the mouse, however, it appears that a gene duplication has given rise to two MBL forms. The MBL-A gene maps to chromosome 14 in a region syntenic to human chromosome 10, whereas the MBL-C gene has been localized to chromosome 19.[28]

The activation of complement upon binding of MBL to carbohydrates on microbial surfaces suggests a role in the innate immune defence. This supposition is supported by the finding that opsonic defects are correlated with MBL deficiency.[29–33]

Evolutionary information often give clues to function and importance and we thus decided to investigate MBL in the chicken.

MATERIALS AND METHODS

Reagents

A chicken liver cDNA library (male broiler-breeders, 7 weeks old, left frontal lobe) in the Lambda ZAP® vector (cat. no. 93 5402) and the helper phage, R408 (cat. no. 20 0252) was purchased from Stratagene (La Jolla, CA). Chicken liver was from a White Leghorn chicken kept at our local animal house. Oligonucleotide primers were synthesized using a DNA synthesizer (381A, Applied Biosystems, Warrington, UK). GeneAmp® Thermostable rTth reverse transcriptase RNA PCR kit was from Perkin Elmer Cetus (Norwalk, CT). TA-cloning kit was from Invitrogen (San Diego, CA). Magic™ Miniprep DNA purification system was from Promega (Madison, WI). Sequenase™ version 2.0 DNA sequencing kit was from USB (United States Biochemicals, Cleveland, OH, cat. no. 70 770). GTG agarose was from FMC, Medinova (Copenhagen, Denmark). The 1 kb DNA-marker was from Gibco (cat. no. 520-5615SA, Gibco BRL, Life Technologies, Gaithersburg, MD). γ-^{32}P-ATP was from Amersham (Amersham, UK). T4-polynucleotide kinase and salmon sperm DNA was from Boehringer Mannheim Biochemica (Mannheim, Germany). PD-10 column was obtained from Pharmacia (Uppsala, Sweden). All other reagents and chemicals were from Sigma (Sigma Chemical Co., St. Louis, MO).

Polymerase chain reaction (PCR) and reverse transcription (RT)–PCR amplification

Based on the published N-terminal partial amino acid sequence of chicken MBL,[9] two degenerate oligonucleotides were synthesized designated cMBL primer 12 (sense primer, 5′-GAYAARCCNGARGARAARATGTA-3′) and cMBL primer 10 (antisense primer, 5′-TCNCCNGGNA RNCCRTTNAC-3′). Chicken liver RNA was purified as described elsewhere.[34] Liver RNA was amplified in a RNA–PCR reaction. After completion of the reverse transcription step another vial was included, containing the same reagents except that the liver RNA was substituted by 1 or 0·1 μl (10^6 and 10^5 phages, respectively) of the cDNA library. The reactions were carried out following the manufacturers recommendations. In brief, reverse transcription for 1 min at 70°, 5 min at 48° and 5 min at 70° using 50 pmol antisense primer (cMBL primer 10), 0.2 μg chicken liver RNA, dNTPs and buffer included in the kit. The reverse transcriptase was inhibited with Mn^{2+} chelating buffer, and the PCR reaction was thereafter carried out by addition of $MgCl_2$ and sense primer (cMBL primer 12). The reactions were carried out for 0·5 min at 94°, 0·5 min at 48° and 0·5 min at 72° for 30 cycles. The obtained PCR products were analysed on a 3% GTG-agarose gel. One μl of the PCR product was mixed with 100 ng pCR™ II-cloning vector, 1 × ligation buffer, T4 DNA ligase and water to a final volume of 10 μl according to the supplier protocol. Ligation took place at 12° for 18 h. Competent *Escherichia coli* (50 μl OneShot™ INV α F′ bacteria) was mixed with 2 μl 0·5 M 2-mercaptoethanol, 1 μl ligation mixture was added, followed by incubation on ice for 30 min, transformation at 42° for 45 s, chilling on ice, addition of 450 μl SOC medium (20g bacto-tryptone, 5g bacto-yeast extract, 0.5 g NaCl, 20 mM glucose per litre of H_2O) and incubation at 37° for 1 h. The transformed cells were spread on Luria–Bertani broth (LB)-plates containing 50 μg ampicillin/ml, and with 5-bromo-4-chloro-3-indolyl-β-D-galactoside and isopropyl-1-thio-β-D-galactopyranoside spread on the surface. White colonies were screened by the PCR reaction described above (omitting the reverse transcription step). Colonies that yielded the 75 bp PCR product were selected, and the plasmids were purified from 10 ml miniprep cultures using the Magic™ Miniprep DNA purification system. Double stranded DNA sequencing was performed on four purified plasmids by the dideoxynucleotide chain termination procedure.

Screening of a chicken liver cDNA library with the perfect match oligonucleotide

A perfect match 30 bp oligonucleotide (5′-GAGA AGATGTATTCCTGTCCCATCATTCAG-3′) was synthesized, based upon the sequence obtained by the procedure described above. The perfect match oligonucleotide was 5′-end labelled in a reaction containing: 50 pmol γ-^{32}P-ATP (2 μl), 50 pmol oligonucleotide (2 μl), 10 × kinase buffer (5 μl), 20 U

T4-polynucleotide kinase (2 µl) and 39 µl ddH_2O at 37° for 30 min. The T4-polynucleotide kinase was inactivated by incubation at 65° for 15 min.

Approximately 5×10^5 phages from the chicken liver cDNA library were plated out, and duplicate filters (cat. no. RPN 132B, Amersham) were lifted, washed in 1·5 M NaCl, 0·5 M NaOH for 2 min, followed by 1·5 M NaCl, 0·5 M Tris, pH 8·0 for 5 min and 30 s in 2 × saline sodium citrate (SSC). The filters were dried and crosslinked using the Stratalinker® UV crosslinker (cat. no. 400071, Stratagene, La Jolla, CA).

The filters were prehybridized for 2 hr at 42° in prehybridization solution (50% v/v formamide, 5 × SSPE, 5 × Denhardt solution, 0·5% (w/v) sodium docecyl sulphate (SDS), 0·1% (w/v) sodium pyrophosphate and 100 µg sonicated salmon sperm DNA per ml.

After purification on a PD-10 column, the labelled probe was added directly to the prehybridization solution, and hybridization was carried out at 42° for 16 h. The filters were washed for 30 min in 4 × SSC, 0·1% SDS at 42°, followed by 30 min in 2 × SSC, 0·1% SDS at 42° and finally in 0·4 × SSC, 0·1% SDS at 65°. The filters were exposed to X-ray film for 3 days between intensifying screens.

Positive plaques were picked, and resuspended in 0·5 ml SM-buffer containing 20 µl chloroform. Specific inserts were confirmed by PCR, as described above, except that the reverse transcription step was omitted.

In vivo *excision of the pBluescript plasmid from the Lambda ZAP vector*

Bacteria (200 µl XL-1 Blue) were superinfected with 200 µl of the isolated vector in SM-buffer and 1 µl of the helper phage R408. After incubation at 37° for 15 min, 5 ml 2 × YT medium were added, and incubation was continued for 3 hr with shaking. The bacteria were killed at 70° for 20 min, and the pBluescript phagemid filamentous phages were separated from the bacteria by spinning at 4000 *g* for 5 min. The supernatant (10 µl) was mixed with 200 µl XL-1 blue host cells, incubated at 37° for 15 min and plated on LB/Amp-plates. Colonies were isolated, and confirmed by PCR, plasmids were isolated from minipreps and sequenced, as described above.

mRNA isolation

Various organs (heart, spleen, liver, kidney and pancreas) were excised immediately after killing the chicken and frozen in liquid nitrogen. The frozen organs were homogenized in a Waring blender in 25 ml of 5·0 M guanidinium thiocyanate (Fluka, Buchs, Switzerland) containing 50 mM sodium citrate pH 7·0, 70 mM β-mercaptoethanol and 0·5% sodium lauryl sarcosine and layered onto 9·5 ml of 5·7 M CsCl, containing 0·1 M ethylenediaminetetra-acetic acid (EDTA)[35] and centrifuged at 140 000 *g* for 26 hr at 18°. After centrifugation the RNA pellet was resuspended in 3 ml ddH_2O, precipitated by addition of 300 µl 2·0 M potassium acetate and 9 ml 99% ethanol and resuspended in 1·5 ml ddH_2O. mRNA was isolated by the PolyATtract® mRNA isolation Systems according to the procedure recommended by the supplier (Promega Corporation, Madison WI). All solutions were pre-treated with 0·1% diethyl pyrocarbonate (DEPC) and filtered through 0·45 µm polysulfone membranes (cat. no. 6781-2504 Watman Scientific Limited, Maidstone, UK).

Northern blotting

For Northern blotting 4 µg mRNA from each organ were adjusted to 100 µl with ddH_2O. Ten µl 2·0 M potassium acetate and 280 µl 99% ethanol were added and the mRNA was precipitated overnight at −20°. After centrifugation, the mRNA pellet was dissolved in 10 µl 50% deionized formamide/2·2 M formaldehyde in MOPS-buffer (20 mM MOPS pH 7·0, 1 mM EDTA, 5 mM sodium acetate) and denatured for 5 min at 67° before running into a 1% formaldehyde–agarose gel. Just before loading 2 µl DNA loading buffer were added to each sample. After running the gel the RNA was blotted onto Hybond N membrane (cat. no. RPN 203N, Amersham) overnight in 20 × SSC, baked for 2 hr at 80°, prehybridized for 2 hr in 6 × SSC, 1% SDS, 10 × Denhardt solution, 10 mg/ml carrier DNA at 67° and hybridized overnight in 6 × SSC, 1% SDS, 10 mg/ml carrier DNA and 8×10^6 c.p.m/ml of radioactive chicken MBL probe at 67°. The chicken MBL probe encodes a 554 bp Pst I–Pvu II fragment of the isolated MBL cDNA clone, containing the CRD domain of the chicken MBL, which was random hexamer primed with ^{32}P-dCTP.[36] The filter was washed two times in 1 × SSC, 0·1% SDS, two times in 0·1 × SSC, 0·1% SDS and finally two times in 0·1 × SSC without SDS for 25 min at 55°. The filter was exposed to X-ray film for 10 days between intensifying screens.

Purification of chicken MBL and peptide sequencing

Peptides DN9, DN12/13, DN22, DN26 and DN29 MBL was purified from serum of White Leghorn chickens as described previously.[22] Briefly, cMBL in delipidized serum (140 ml) was absorbed on the cells of Ra chemotype strain (rfb388) of *Salmonella* serovar Typhimurium, eluted with *N*-acetylglucosamine and finally passed through an anti-chicken immunoglobulin G (IgG) column (Bethyl Laboratories, Inc., Tokyo, Japan). The effluent was reduced and *S*-carboxymethylated as described.[37,38] The alkylated protein solution was loaded on an Asahipak C4P-50 column (0·46 × 15 cm, Asahi Biochemicals, Tokyo, Japan) and eluted with a linear gradient of acetonitrile in 0·1% trifluoroacetic acid (60 min at 40° at a flow rate of 0·8 ml/min). Fractions were collected, analysed by SDS–Polyacrylamide gel electrophoresis (PAGE) and a fraction containing the 33 000 MW and 34 000 MW polypeptides was subjected to peptide sequencing (these two bands could not be separated). The polypeptides were digested for 18 hr at 37° with endoprotease ASP-N (Boehringer Mannheim) at an enzyme/substrate ratio of 1 : 80 in 50 mM phosphate buffer containing 1 M urea, pH 8·0. Proteolytic fragments were separated by reverse phase high pressure liquid chromatography (HPLC) on a TSKgel ODS-120T column (0·46 × 25 cm, Toyo Pearl, Tokyo, Japan) using a linear gradient of acetonitrile in 0·1% trifluoroacetic acid (80 min at 40° at a flowrate of 0·8 ml/min). Fractions DN9, DN12/13, DN22, DN26 and DN29 were isolated. By further fractionation, two subfractions DN22a and DN22b were obtained from fraction DN22, and five subfractions DN26a to DN26e were obtained from the fraction DN26.

Amino acid sequences of these fractions were determined by step-wise Edman degradation using an automated protein sequencer (Applied Biosystems, Model 470/120A).

Peptides ACW-1 and ACW-3 cMBL was purified as described elsewhere.[9] Calcium-dependent serum lectins from

White Leghorn chickens were purified by two sequential affinity chromatographies on TSK-beads derivatized with ManNAc and elution with EDTA and ManNAc, respectively. Chicken anti-carbohydrate antibodies were removed with anti-chicken immunoglobulin beads and the MBL further purified by anion exchange chromatography and gel permeation chomatography. SDS–PAGE analysis under reducing conditions revealed only two bands with molecular weights of 33 000 and 34 000 MW.[9] The preparation was digested with immobilized collagenase as described previously.[9] Twenty micrograms intact or 20 μg collagenase treated protein solution was run on SDS–PAGE under reducing conditions and the proteins transferred to a polyvinylidene diflouride membrane (Immobilon, Millipore, Bedford, MA). The blot was stained with Ponceau S, the two 33 000 and 34 000 MW bands (untreated) or the 21 000 MW band (collagenase treated) were cut out and sequenced on an Applied Biosystem 470A/120A sequencer.[39]

Purification of cMBL from egg yolk was done as described for peptides ACW-1 and ACW-3, except that 20 egg yolks were used instead of the chicken serum.

Peptide ACW-2 (liver) Livers from 10 White Leghorn chickens were homogenized at 0° in 200 ml buffer composed of 50 mM Tris/HCl, 100 mM NaCl, 5 mM iodoacetamide, 5 mM cyclocapron (Kabi, Stockholm, Sweden), 5 mM EDTA, 10 U Trasylol/ml, 1 mM phenyl-methyl–sulfonyl-fluoride (PMSF) and 1% (v/v) Emulfogene. After centrifugation (10^4 ***g***, 30 min, 4°), the supernatant (200 ml) was diluted with the above mentioned column buffer and passed through a precolumn (150 ml Sepharose 4B, Pharmacia) and through a column of mannan–Sepharose (100 ml). The mannan–Sepharose column was washed with column buffer until the absorption at 280 nm was below 0·05, and eluted with column buffer containing 10 mM EDTA instead of $CaCl_2$. Fractions containing proteins were pooled, recalcified to 25 mM $CaCl_2$ and pH readjusted to 7·4. The pool was passed through a new mannan–Sepharose column (50 ml), which after washing was eluted first with column buffer containing 50 mM ManNAc and then with column buffer containing 10 mM EDTA instead of $CaCl_2$.

Western blotting

Samples were reduced and denatured by heating at 100° in 40 mM dithiothreitol (DDT), 1·5% (w/v) SDS, 5% (v/v) glycerol, 0·1 M Tris, pH 8·0, and alkylated by the addition of iodoacetamide to 90 mM. Unreduced samples were prepared by omitting dithiothreitol. The samples were run on 4–20% polyacrylamide gradient gels using the Laemmli discontinuous buffer system.[40] The molecular weight standards were Rainbow® coloured protein molecular weight markers (code no. RPN 755, Amersham). The gel was blotted onto a polyvinylidene difluoride (PVDF) membrane (Immobilon, Millipore). The Western blot was blocked in Tris-buffered saline (TBS) containing 0·1% Tween 20 for 15 min followed by incubation with a monoclonal anti-chicken MBL antibody (HYB 182-2, the State Serum Institute, Copenhagen, Denmark) at 0·3 μg per ml TBS/Tween. The blot was washed trice in TBS/Tween and incubated with alkaline-phosphatase-conjugated rabbit anti-mouse immunoglobulin (cat. no. D0314, Dakopatts, Glostrup, Denmark) diluted 1/5000 in TBS/Tween. After additional washing, the blot was developed with 100 μg nitroblue tetrazolium and 50 μg potassium-5-bromo-4-chloro-3-indolylphosphate per ml, 2 mM $MgCl_2$, 0·1 M ethanolamine, pH 9·0.

Phylogenetic analysis

The deduced amino acid sequence of chicken MBL was compared with the sequences of human MBL,[26] mouse MBL-A and MBL-C,[7] rat MBL-A and MBL-C,[14] rhesus MBL-A and MBL-C[2] and bovine MBL using a maximum likelihood method for calculating distances and the neighbour-joining algorithm for tree construction as included in the PHYLIP software package (Felsenstein, J., Neighbor-Joining/UPGMA method version 3.51c).

RESULTS

Based on our published N-terminal sequence of chicken MBL,[9] two degenerate oligonucleotides were synthesized. Using these primers in a RT–PCR reaction on chicken liver RNA, a band of approximately 75 bp was amplified (Fig. 1, lane 1). A band of the same size was obtained using the same primers directly on two different concentrations of a chicken liver cDNA library in a PCR reaction without the initial reverse transcription step (Fig. 1, lanes 3 and 4). The controls using the same reagents omitting template RNA or DNA were negative. The PCR fragment was cloned into the pCR® II-cloning vector and the 75 bp insert was sequenced. The deduced amino acid sequence matched perfectly with the N-terminal sequence of the purified protein. A perfect match cDNA probe was synthesized, radiolabelled and used to screen the chicken liver cDNA library. Eight positive clones were isolated and the presence of a specific insert was confirmed by PCR. After rescreening twice, the phagemids were excised using the helper phage R408. The eight plasmids all showed the same size on gel electrophoresis (1·7 kb, data not shown). The inserts in the plasmids were sequenced partially and shown to contain the same full length cDNA insert, whereafter the whole inserts in the plasmids were sequenced in both directions.

The amino acid sequences obtained by stepwise Edman degradation of intact cMBL and proteolytic fragments are shown in Fig. 2. All sequences except ACW-2 (liver) were obtained from cMBL purified from serum.

The open reading frame of 714 bp of the cMBL encoding cDNA sequence is shown in Fig. 3. The cDNA contains 15 nucleotides corresponding to a leader peptide of five residues, followed by 699 nucleotides corresponding to the coding

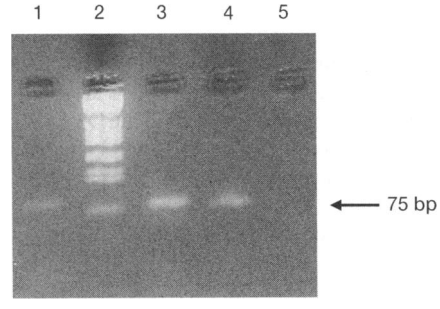

Figure 1. RT–PCR on chicken liver RNA (lane 1) and PCR on chicken liver cDNA library (10^6 phages, lane 3 and 10^5 phages, lane 4). Lane 2, molecular weight marker and lane 5, control without template DNA/RNA.

```
ACW-1          L L T X D K P E E K M Y S C P I I Q C S A P A V N G L hP G

ACW-2(LIVER)   G E hK G Q hK G E R G I V V T D D L H R Q I T D L E A K I

ACW-3          G E R G I V V T D D L H R Q I T D L E A K I R V L E D D L
               S R Y K K A L

DN9            D S G K W N

DN12/13        D D L H R Q I T

DN22-a         D G P hK G E H G D P G E G L R G L Q G L hP G K A G P Q G L
               L E V G P Q G E X G Q Q G

DN22-b         D D L S R Y K K A L S L K

DN26-a         D L E A K I R V L E

DN26-c         D L V D P S S Q A Y I G I S

DN26-e         D V V N V G K K M F V S T G K K Y N F E K G K S L C A K A
               G S V L A S P X N E A

DN29           D A Q T E G R F M Y L S G G P L T Y S N W K P G E P N N H
               K N E
```

Figure 2. Amino acid sequence of either intact chicken MBL (ACW-1) or proteolytic fragments obtained by peptide sequencing. All peptides except ACW-2 (liver) were obtained from chicken MBL purified from serum, whereas ACW-2 (liver) was obtained from MBL purified from chicken liver.

sequence of the mature protein of 233 residues. The N-terminus of the mature protein is known from previous amino acid sequencing.[9]. The sequence does not present any site for N-type glycosylation. Assuming no posttranslational modifications, the molecular weight of the protein, is 25 648 g/mol with an isoelectric point of 5·93 and an absorption coefficient of 13·0 ($A_{280\ nm}$ 1%, 1 cm).

There are four discrepancies between the deduced sequence and the sequence of the aligned peptides (Fig. 3) at amino acid position no. 40 (K, H), no. 63 (K, L), no. 119 (I, V), no. 163 (I, V). The HPLC identified five hydroxylated residues, at position no. 28, 37, 54, 72 and 75. At position 75 the deduced amino acid is lysine, the ACW-2 shows hydroxylated lysine and the peptide DN-22a showed glutamine.

Northern blotting analysis (Fig. 4), using a cMBL specific probe encoding the neck region and the whole CRD, showed a 1·7 kb cMBL transcript in the liver only whereas other organs tested were negative (heart, spleen, kidney and pancreas).

In Fig. 5 the deduced amino acid sequence is aligned with the sequences of mammalian MBLs. The overall sequence homologies are 40% with human MBL, 38% with mouse MBL-A, 39% with mouse MBL-C, 38% with rat MBL-A, 37% with rat MBL-C, 40% with rhesus monkey MBL-A and 40% with rhesus monkey MBL-C. The N-terminal region of chicken MBL consists of 25 amino acids followed by 56 residues containing 19 Gly-X-Y repeats. The seventh repeat (G-E-G) represents an interruption such as those found in mammalian MBLs.[2,41] The neck region consists of 38 amino acids which is two to four residues larger than mammalian MBLs. The CRD consists of 114 residues with 34 residue identities between chicken and mammalian MBLs. Between mammalian MBLs there are 47 conserved residues in this region. The 14 conserved residues and 18 residues conserved in character previously reported in all C-type lectins are also found in chicken MBL.[42]

Figure 6 shows a comparison by Western blotting analysis of MBL purified from chicken serum, yolk and liver. Serum and yolk show identical MBL, with a double band at 33 000 and 34 000 MW under reducing conditions (lanes 2 and 3). This double band has previously been shown to contain polypeptides with different pI distributions, but identical N-terminal sequences (30 amino acids sequenced). Collagenase digestion of these two bands yielded one band at 18 000 MW when unreduced and 21 000 MW when reduced, indicating that the difference between the two polypeptides resides in the CRD region, and is possibly due to disulphide bonds.[9] Analysing these two preparations unreduced reveals a ladder of bands (lanes 6 and 7) with molecular weight about 70 000 MW and above. The lectin purified from liver homogenates run under non-reducing conditions showed a prominent 27 000 MW band, no band at 70 000 MW, but several high molecular weight bands (lane 5). Under reducing conditions all the MBL is found at 30 000 MW (lane 1). However, elution with EDTA after the ManNAc elution, revealed MBL with the same mobility as serum MBL (lane 4). Small amounts of 27 000 MW MBL is seen together with high molecular weight MBL (lane 8). The 30 000 MW band did not yield an N-terminal sequence presumably due to N-terminal blockage. Peptides recovered after digestion with collagenase revealed the sequence shown in Fig. 2 (ACW-2). Upon fractionation of the liver lectin by size-permeation chromatography on a Superose 6 column, as described previously for serum cMBL,[9] the lectin eluted with ManNAc (Fig. 6, lanes 1 and 5) eluted in a symmetrical peak at 330 000 MW, whereas the lectin eluted with EDTA (Fig. 6, lanes 4 and 8) gave two symmetrical peaks at 330 000 MW and 750 000 MW (data not shown). Only the peak at 750 000 MW has been observed for MBL purified from serum.[9]

Evolutionary analysis (Fig. 7) indicates, that bovine and chicken MBL have no preferential association with either the A or C type clades of primate and rodent MBLs. Thus, the nodes joining the avian branch and the tree of mammalian

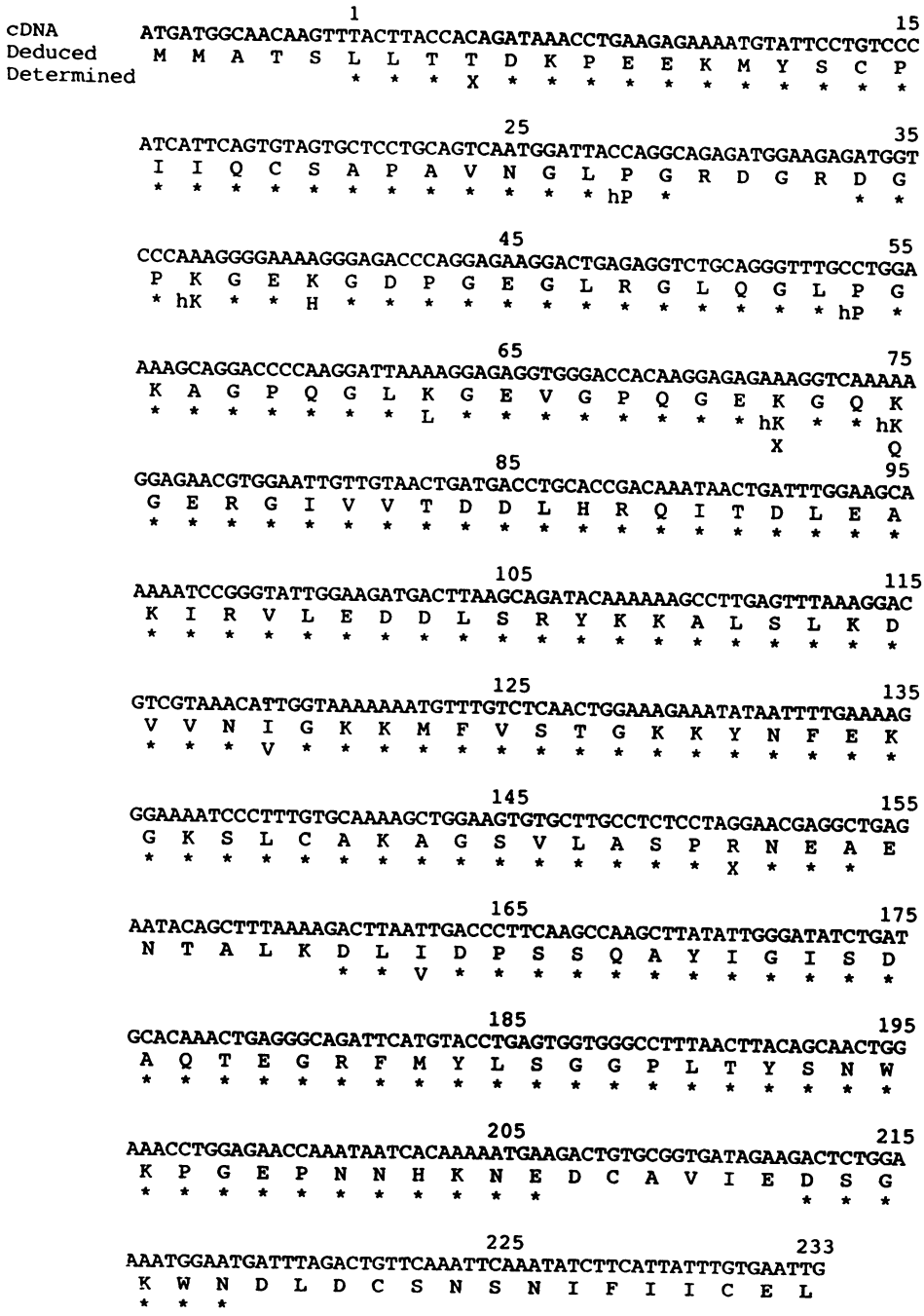

Figure 3. Alignment of the cDNA sequence with the deduced amino acid sequence and sequences obtained from peptides shown in Fig. 2. Asterix denotes identical residues; X, unidentified residue; h, hydroxylated amino acid.

species might represent an undiversified ancestor of MBL-A and C at the level of the Cotylosauria. It must be emphasised that bootstrap values below 100 only occurred at the nodes joining the bovine (bootstrap value 43) and avian (bootstrap value 68) MBL branches to the tree.

DISCUSSION

We have cloned and sequenced a chicken serum lectin, which appears to be the avian analogue to mammalian MBL. The deduced amino acid sequence, except for 36 amino acid residues, was also determined by peptide sequencing. The amino acid sequence represents a mature protein of 233 residues and a leader sequence of five residues. The amino acid sequence deduced from the cDNA clone differs from those obtained by peptide sequencing at four residues. At position 40, nucleotide sequencing yields lysine and peptide sequencing histidine, at position no. 63, lysine were found leucine, at position 119 isoleucine and valine, and at position 163 isoleucine and valine. The discrepancies at positions 40 and 63 would both require two basepair mutations and are thus possibly explained by peptide sequencing errors. Lysine

Figure 4. Northern blotting analysis of poly A^+ RNA from different chicken organs. Five microlitres poly A^+ RNA from the indicated organs were run in a formaldehyde gel, blotted onto filter and hybridized with a 554 bp Pst I–Pvu II fragment of the isolated MBL cDNA clone. The molecular size of the band was calculated using 28S and 18S RNA bands as molecular size markers.

(40) is encoded by the codon AAG and this is unlikely to represent a mutation (two basepair exchange) or misinterpretation of nucleic acid sequences. Lysine (63) is encoded by the codon AAA and leucine can potentially be encoded by TTG, TTA, CTT, CTC, CTG or CTA, so this exchange also requires at least two mutations. The other discrepancies requiring only one basepair substitution are likely to be caused by different genotypes of the chicken strains. Isoleucines 119 and 163 are encoded by the codon ATT, and mutation to GTT would give rise to valine. A discrepancy between the previously published N-terminal sequence[9] and the deduced sequence is found at residue 4, where the published sequence mistakenly showed a cysteine residue. Amino acid 75 (lysine) of liver MBL was found to be lysine, but MBL purified from serum gave glutamine. It must be emphasized that the cDNA was sequenced in both directions and no ambiguities were found.

When comparing the known MBL sequences (Fig. 5) the most striking difference is the nine amino acid shorter collagenous region in the rodent MBL-A forms. Whether this stretch is involved in the activation of the complement system remains to be elucidated. It has been shown, that murine MBL-A is able to activate complement, but on the other hand it has been shown, that recombinant human MBL lacking these nine amino acids was unable to activate complement.[16,43]

In chicken serum MBL, there are only two cysteines in the N-terminal region, whereas the other MBLs, except mouse and rat MBL-C have three cysteines. These cysteines are thought to be involved in both intra- and interchain disulphide bonds and the results in the chicken indicate that only two cysteines are required for forming the bouquet like structure of 750 000 MW.[9] At the position of the seventh Gly-X-Y repeat an interrupt is found with one amino acid missing. The same interrupt has been found in this very conserved region in all mammalian MBL sequences. SPA, another collectin also contains an interruption, thought to be responsible for the kink in the tertiary structure of the collagen-like triple helices observed under electron microscopy.[44]

Chicken MBL shows two extra amino acids between the collagen-like region and the neck region. In the neck region one finds two extra amino acids in chicken MBL, rodent and rhesus monkey MBL-A, as compared to human MBL and rodent/rhesus MBL-C. It is interesting to note that the neck region with only one conserved leucine residue is the least conserved region of the MBLs, but the hydrophobic nature of several residues are highly conserved, which form the triple helical coiled coil.[13,45,46]

When comparing the chicken MBL sequence with mammalian MBLs, 35 identical residues out of 114 (31%) can be found in the CRD-domain whereas in the collagenous region 32 amino acids out of 56 (57%) are conserved. The 14 completely conserved residues and 18 residues conserved in character found in all C-type lectins are also conserved in chicken MBL.[42] No site for N-type glycosylation (N-X-T or N-X-S) is found in the sequence, which contradicts the previously reported experiments indicating the presence of one or more N-linked sialyated oligosaccharides.[9] A possible explanation of this discrepancy could be glycosylation other than N- or O-glycosidic bonds to amino acids.[47] We are presently unable to explain this discrepancy.

Northern blotting was carried out with a chicken MBL specific probe spanning the 'neck region' and the CRD. A transcript (1·7 kb) was found only in liver (Fig. 4). Using the whole cloned cDNA as probe slightly larger transcripts (1·8 kb) could be seen in kidney, pancreas and heart tissue. This is probably caused by cross-hybridization of the probe with transcripts encoding collagen or collagen-like proteins.

Lectin was purified from chicken liver extracts and in Western blotting analysis using a monoclonal anti-chicken MBL antibody staining of 30 000, 33 000 and 34 000 MW polypeptides was observed. The two latter are likely to represent the same MBL polypeptides as seen when analysing serum MBL which have been shown to be derived from the high molecular weight, intact cMBL molecule of 750 000 MW.[9] The 30 000 MW polypeptide contains intra-chain disulphide bonds and represents the polypeptides of a protein with an intact size of 330 000 MW. It was however not possible to get a N-terminal sequence of the protein probably because it was N-terminally blocked. Sequencing of peptides obtained by collagenase treatment showed identical amino acids to the serum form (28 amino acids sequenced). This sequence covers part of the collagen-like sequence and 16 amino acids identical with the neck region of chicken MBL purified from serum (Fig. 2). Comparison of the neck region in different mammals reveals, that this region is not conserved between the A- and C-forms. It therefore seems unlikely, that the 30 000 MW polypeptide should represent an avian analogue to the mammalian MBL-C, although this region could for some reason be more conserved in birds than in mammals. It seems more likely that liver cells contains an immature form of MBL, which is able to bind to mannan, and that lack of glycosylation and hydroxylation may account for the 3000–4000 MW smaller size observed on SDS–PAGE. In humans, liver has

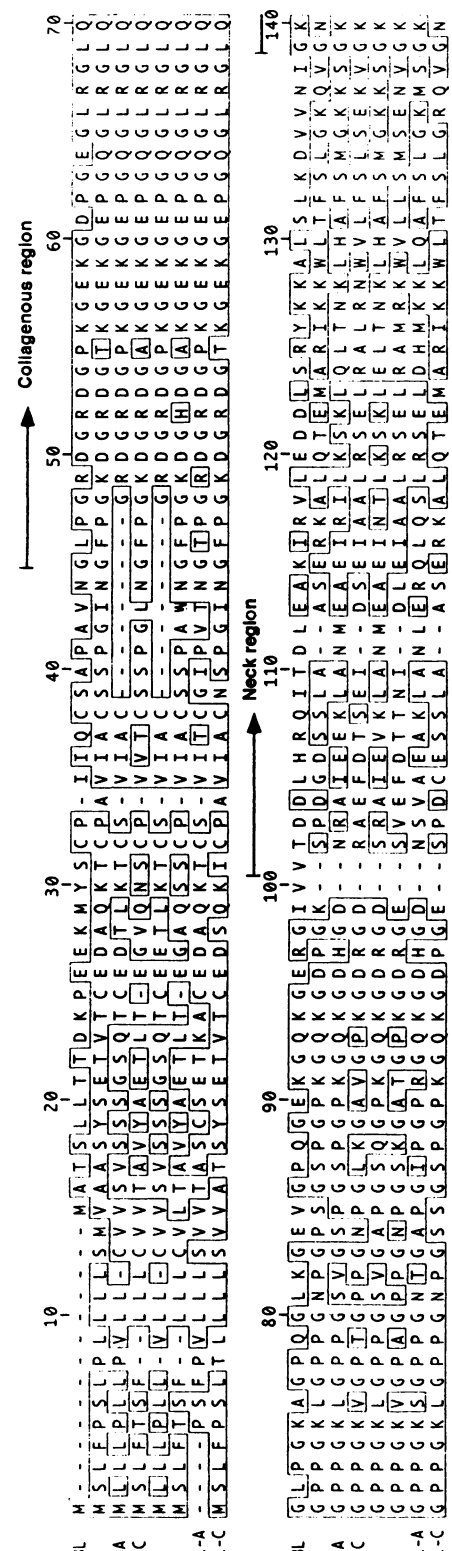

Figure 5. Alignment of the deduced chicken MBL sequence with man MBL, rodent and rhesus monkey MBL-A and MBL-C. Identical residues are indicated by a box, and the different regions of the proteins denoted above the sequences. Asterix denotes the conserved residues in all C-type lectins identified.

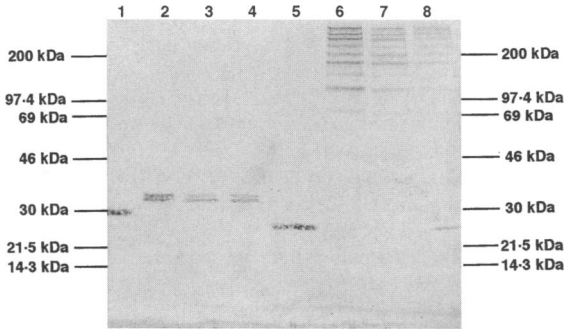

Figure 6. Western blotting analysis of chicken MBL isolated from liver (lanes 1, 5; ManNAc-eluate and lanes 4, 8; EDTA eluate), serum (lanes 2, 6) or yolk (lanes 3, 7). Lanes 1–4 show fractions analyzed under reducing conditions and lanes 5–8 unreduced.

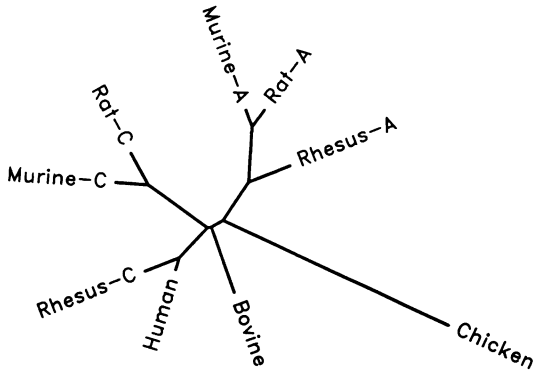

Figure 7. Unrooted phylogeny of mammalian and avian MBLs using a maximum likelihood method for calculating distances and the neighbour-joining method for tree construction.

been found to contain a form of MBL not found in serum, although both are composed of the same polypeptides.[16] The liver form seems to be composed of only nine polypeptides whereas 18 participate in forming the serum MBL. Only the serum form has the ability to activate the complement system.[16] This situation seems similar in chickens, because the liver contains MBL molecules of 330 000 MW and 750 000 MW whereas in serum only the 750 000 MW form is observed, and it is only the serum form that can activate the complement system (unpublished observation).

NOTES

The cDNA sequence has been submitted to GenBank with the accession number AF 02 2226.

REFERENCES

1. Kawasaki N., Kawasaki T. & Yamashina I. (1983) Isolation and characterization of a mannan-binding protein from human serum. *J Biochem Tokyo* **94,** 937.
2. Mogues T., Ota T., Tauber A.I. & Sastry K.N. (1996) Characterization of two mannose-binding protein cDNAs from rhesus monkey (*Macaca mulatta*): structure and evolutionary implications. *Glycobiology* **6**(5), 543.
3. Kozutsumi Y., Kawasaki T. & Yamashina I. (1980) Isolation and characterization of a mannan-binding protein from rabbit serum. *Biochem Biophys Res Commun* **95,** 658.
4. Holmskov U., Holt P., Reid K.B.M., Willis A.C., Teisner B. & Jensenius J.C. (1993) Purification and characterization of bovine mannan-binding protein. *Glycobiology* **3,** 147.
5. Storgaard P., Nielsen E.H., Andersen O. *et al.* (1996) Isolation and characterization of porcine mannan-binding proteins of different size and ultrastructure. *Scand J Immunol* **43,** 289.
6. Oka S., Ikeda K., Kawasaki T. & Yamashina I. (1988) Isolation and characterization of two distinct mannan-binding proteins from rat serum. *Arch Biochem Biophys* **260,** 257.
7. Sastry K., Zahedi K., Lelias J.M., Whitehead A.S. & Ezekowitz R.A. (1991) Molecular characterization of the mouse mannose-binding proteins. The mannose-binding protein A but not C is an acute phase reactant. *J Immunol* **147,** 692.
8. Sugi S. & Hirora Y. (1994) Identification and carbohydrate specificity of a chicken serum mannan-binding protein reactive with Ra chemotype strain of *Salmonella typhimurium. J Vet Med Sci* **56,** 747.
9. Laursen S.B., Hedemand J.E., Thiel S. *et al.* (1995) Collectin in a non-mammalian species: Isolation and characterization of mannan-binding protein (MBP) from chicken serum. *Glycobiology* **5,** 553.
10. Malhotra R., Thiel S., Jensenius J.C. & Sim R.B. (1991) Interaction of Lung surfactant protein A with C1q receptor. *Complement Inflammation* **8,** 188.
11. Drickamer K. (1988) Two distinct classes of carbohydrate-recognition domains in animal lectins. *J Biol Chem* **263,** 9557.
12. Weis W.I., Kahn R., Fourme R., Drickamer K. & Hendrickson W.A. (1991) Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science* **254,** 1608.
13. Hoppe H.J. & Reid K.B.M. (1994) Trimeric C-type lectin domains in host defence – recent crystal structures, established for fragments of human and rat mannose-binding proteins, indicate that the triple-stranded alpha-helical coiled coil present in these collectins is responsible for the trimeric orientation of C-type lectin domains. *Structure* **2,** 1129.
14. Drickamer K., Dordal M.S. & Reynolds L. (1986) Mannose-binding proteins isolated from rat liver contain carbohydrate-recognition domains linked to collagenous tails. Complete primary structures and homology with pulmonary surfactant apoprotein. *J Biol Chem* **261,** 6878.
15. Wild J., Robinson D. & Winchester B. (1983) Isolation of mannose-binding proteins from human and rat liver. *Biochem J* **210,** 167.
16. Kurata H., Sannoh T., Kozutsumi Y., Yokota Y. & Kawasaki T. (1994) Structure and function of mannan-binding proteins isolated from human liver and serum. *J Biochem (Tokyo)* **115,** 1148.
17. Thiel S., Holmskov U., Hviid L., Laursen S.B. & Jensenius J.C. (1992) The concentration of the C-type lectin, mannan-binding protein, in human plasma increases during an acute phase response. *Clin Exp Immunol* **90,** 31.
18. Ezekowitz R.A., Day L.E. & Herman G.A. (1988) A human mannose-binding protein is an acute-phase reactant that shares sequence homology with other vertebrate lectins. *J Exp Med* **167,** 1034.
19. Arai T., Tabona P. & Summerfield J.A. (1993) Human mannose-binding protein gene is regulated by interleukins, dexamethasone and heat shock. *Quart J Med* **86,** 575.
20. Lu J.H., Thiel S., Wiedemann H., Timpl R. & Reid K.B. (1990) Binding of the pentamer/hexamer forms of mannan-binding protein to zymosan activates the proenzyme C1r2C1s2 complex, of the classical pathway of complement, without involvement of C1q. *J Immunol* **144,** 2287.

21. IKEDA K., SANNOH T., KAWASAKI N., KAWASAKI T. & YAMASHINA I. (1987) Serum lectin with known structure activates complement through the classical pathway. *J Biol Chem* **262,** 7451.
22. JI Y.H., FUJITA T., HATSUSE H., TAKAHASHI A., MATSUSHITA M. & KAWAKAMI M. (1993) Activation of the C4 and C2 components of complement by a proteinase in serum bactericidal factor, Ra reactive factor. *J Immunol* **150,** 571.
23. DRICKAMER K. & MCCREARY V. (1987) Exon structure of a mannose-binding protein gene reflects its evolutionary relationship to the asialoglycoprotein receptor and nonfibrillar collagens. *J Biol Chem* **262,** 2582.
24. TAYLOR M.E., BRICKELL P.M., CRAIG R.K. & SUMMERFIELD J.A. (1989) Structure and evolutionary origin of the gene encoding a human serum mannose-binding protein. *Biochem J* **262,** 763.
25. SASTRY R., WANG J.S., BROWN D.C., EZEKOWITZ R.A.B., TAUBER A.I. & SASTRY K.N. (1995) Characterization of murine mannose-binding protein genes Mbl1 and Mbl2 reveals features common to other collectin genes. *Mamm Genome* **6,** 103.
26. SASTRY K., HERMAN G.A., DAY L. *et al.* (1989) The human mannose-binding protein gene. Exon structure reveals its evolutionary relationship to a human pulmonary surfactant gene and localization to chromosome 10. *J Exp Med* **170,** 1175.
27. KÖLBLE K., LU J., MOLE S.E., KALUZ S. & REID K.B.M. (1993) Assignment of the human pulmonary surfactant protein D gene (SFTP4) to 10q22–q23 close to the surfactant protein A gene cluster. *Genomics* **17,** 294.
28. WHITE R.A., DOWLER L.L., ADKISON L.R., EZEKOWITZ R.A.B. & SASTRY K.N. (1994) The murine mannose-binding protein genes (Mbl 1 and Mbl 2) localize to Chromosomes 14 and 19. *Mamm Genome* **5,** 807.
29. SUPER M., THIEL S., LU J., LEVINSKY R.J. & TURNER M.W. (1989) Association of low levels of mannan-binding protein with a common defect of opsonisation. *Lancet* **2,** 1236.
30. SUMIYA M., SUPER M., TABONA P. *et al.* (1991) Molecular basis of opsonic defect in immunodeficient children. *Lancet* **337,** 1569.
31. LIPSCOMBE R.J., LAU Y.L., LEVINSKY R.J., SUMIYA M., SUMMERFIELD J.A. & TURNER M.W. (1992) Identical point mutation leading to low levels of mannose binding protein and poor C3b mediated opsonisation in Chinese and Caucasian populations. *Immunol Lett* **32,** 253.
32. SUMMERFIELD J.A., RYDER S., SUMIYA M. *et al.* (1995) Mannose binding protein gene mutations associated with unusual and severe infections in adults. *Lancet* **345,** 886.
33. GARRED P., MADSEN H.O., KURTZHALS J.A.L. *et al.* (1992) Diallelic polymorphism may explain variations of the blood concentration of mannan-binding protein in Eskimos, but not in black Africans. *Eur J Immunogene* **19,** 403.
34. CHORNAZYNSKI P. & SACCHI N. (1987) Single-step method of RNA isolation by acid guaridium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162,** 156.
35. GLISIN V., CRKVENJAKOV R. & BYUS C. (1974) Ribonucleic acid isolation by cesium chloride centrifugation. *Biochemistry* **13,** 2633.
36. FEINBERG A.P. & VOGELSTEIN B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132,** 6.
37. GURD F.R.W. (1967) Carboxymethylation. *Methods Enzymol* **11,** 424.
38. HIRS C.H.W. (1967) Reduction and S-carboxymethylation of proteins. *Methods Enzymol* **11,** 199.
39. CRAWFORD C., BROWN N.R. & WILLIS A.C. (1990) Investigation of the structural basis of the interaction of calpain II with phospholipid and with carbohydrate. *Biochem J* **265,** 575.
40. LAEMMLI U.K. (1970) Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. *Nature* **227,** 680.
41. HOPPE H.J. & REID K.B.M. (1994) Collectins –soluble proteins containing collagenous regions and lectin domains – and their roles in innate immunity. *Protein Sci* **3,** 1143.
42. DRICKAMER K. (1993) Ca^{2+}-dependent carbohydrate-recognition domains in animal proteins. *Curr Opin Struct Biol* **3,** 393.
43. HOLT P., HOLMSKOV U., THIEL S., TEISNER B., HOJRUP P. & JENSENIUS J.C. (1994) Purification and characterization of mannan-binding protein from mouse serum. *Scand J Immunol* **39,** 202.
44. VOSS T., EISTETTER H., SCHAFER K.P. & ENGEL J. (1988) Macromolecular organization of natural and recombinant lung surfactant protein SP 28–36. Structural homology with the complement factor C1q. *J Mol Biol* **201,** 219.
45. HOPPE H.J., BARLOW P.N. & REID K.B.M. (1994) A parallel three stranded helical bundle at the nucleation site of collagen triple-helix formation. *FEBS Lett* **344,** 191.
46. SHERIFF S., CHANG C.Y.Y. & EZEKOWITZ R.A.B. (1994) Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil. *Nature Struct Biology* **1,** 789.
47. HOFSTEENGE J., MÜLLER D.R., DE BEER T., LÖFFLER A., RICHTER W.J. & VLIEGENTHART J.F.G. (1994) New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase U_s. *Biochemistry* **33,** 13 524.