Characterization of the Oligomer Structure of Recombinant Human Mannan-binding Lectin*

Received for publication, November 4, 2004, and in revised form, January 14, 2005
Published, JBC Papers in Press, January 14, 2005, DOI 10.1074/jbc.M412472200

Pia H. Jensen‡§, Dietmar Weilguny‡, Finn Matthesien‡, Kirsten A. McGuire‡, Lei Shi§, and Peter Højrup§

From the ‡Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark, §NatImmune A/S, Fruebjergvej 3, Box 3, DK-2100 Copenhagen, Denmark, and ¶Laboratory of Developmental Immunology, GRUJ1402, MassGeneral Hospital for Children, Harvard Medical School, Boston, Massachusetts 02114

Mannan-binding lectin (MBL) belongs to a family of proteins called the collectins, which show large differences in their ultrastructures. These differences are believed to be determined by different N-terminal disulfide-bonding patterns. So far only the bonding pattern of two of the simple forms (recombinant rat MBL-C and bovine CL-43) have been determined. Recombinant MBL expressed in human cells was purified, and the structure of the N-terminal region was determined. Preliminary results on human plasma-derived MBL revealed high similarity to the recombinant protein. Here we report the structure of the N-terminal part of recombinant human MBL and present a model to explain the oligomerization pattern. Using a strategy of consecutive enzymatic digestions and matrix-assisted laser desorption ionization mass spectrometry, we succeeded in identifying a number of disulfide-linked peptides from the N-terminal cysteine-rich region. Based on these building blocks, we propose a model that can explain the various oligomeric forms found in purified MBL preparations. Furthermore, the model was challenged by the production of cysteine to serine mutants of the three N-terminally situated cysteines. The oligomerization patterns of these mutants support the proposed model. The model indicates that the polypeptide dimer is the basic unit in the oligomerization.

Mannan-binding lectin (MBL) is a serum protein that acts in innate immunity. It is a C-type lectin that recognizes and binds to specific sugars such as D-mannose and N-acetylgalactosamine present on pathogen surfaces (1). The main functions in innate immunity are 1) the opsonin effect, where it binds to the surface of the pathogen and thereby enhances its clearance from the bloodstream and 2) the ability to activate the complement cascade via the lectin pathway. Activation of the complement cascade requires binding of the MBL-associated serine proteases to the oligomeric forms of MBL (3–5) and leads to the formation of the membrane attack complex that perforates the cell membrane of the pathogen. The function of MBL is thus dependent on its oligomeric structure, since the small oligomer forms act as opsonins and the large oligomer forms activate complement (6–8).

The overall polypeptide structure of MBL is similar to that of the other collectins (surfactant protein A, surfactant protein D, conglutinin, CL-43, liver collectin 1, and CL-46). It includes a short, cysteine-rich N-terminal stretch (aa 1–21), a collagen-like region (aa 22–81) with one interruption (aa 43–44) that causes the collagen-like structure to bend, a neck region (aa 82–115), and a carbohydrate recognition domain (aa 116–228) (9). This domain confers the carbohydrate specificity of MBL and is stabilized by two disulfide bonds (1). Due to the collagen-like domain, MBL forms homotrimers, designated the MBL subunit. The collagen-like structure is stabilized by the presence of hydroxyprolines and glycosylated hydroxylsines (10). The subunit structures assemble from the C to the N terminus. The neck region initiates the folding (11), and the collagen-like region zips toward the N terminus, creating trimeric subunits. The structure is finally stabilized by intrasubunit disulfide bonds in the N-terminal region (12). The oligomer structure of MBL is similar to the structure of C1q, the primary component of the classical pathway of complement (13), where the bouquet-like forms arise from the formation of intersubunit disulfide bonds in the N-terminal region (14). However, whereas C1q is made up of three different polypeptide chains, MBL consists of homotrimers. In addition, C1q only exists as hexamers, whereas MBL exists as anything from dimers to octamers. The disulfide-bonding pattern of C1q thus cannot be expected to be identical to that of MBL.

The elucidation of the structure of MBL is complicated by the fact that the polypeptide chain of MBL is very heterogeneous. In addition to several post-translational modifications, there are three well documented mutations in the collagen-like region (15–17). These mutations all lead to amino acid substitutions, which distort the collagen-like region and inhibit the correct formation of the oligomer forms of MBL. In addition to the heterogeneity of the polypeptide chain, promoter polymorphisms (18) result in highly variable amounts of MBL in the blood.

The objective of the present work was to map the disulfide-bonding pattern of the N-terminal part of MBL. The work was performed using recombinant protein due to the availability of larger amounts and the fact that the recombinant protein would be more homogenous than MBL purified from a pool of plasma.

This paper is available on line at http://www.jbc.org

1 The abbreviations used are: MBL, mannan-binding lectin; rMBL, recombinant MBL; rMBL, recombinant human MBL; pMBL, plasma-derived MBL; aa, amino acid; MALDI, matrix-assisted laser desorption/ ionization; MS, mass spectrometry; CL-43 and -46, collectin 43 and 46, respectively; IAA, iodoacetamide; Q-TOF, quadrupole time-of-flight; DTT, dithiothreitol; HPLC, high pressure liquid chromatography.
**EXPERIMENTAL PROCEDURES**

**Materials**—α-Cyano-4-hydroxycinnamic acid was from Aldrich (re-crystallized in boiling acetonitrile). Porcine trypsin was a gift from Novo Nordisk A/S. 3–8% NuPAGE gradient gels and 45-nm polyvinylidene difluoride membranes were from Novex. Primary anti-human MBL antibody HYB151-01 was purchased from The Antibody Shop (The State Serum Institute), and polyclonal horseradish peroxidase anti-mouse antibody was from DAKO. The SuperSignal West Pico substrate was from Pierce. The XCell SureLock & blot M Mini Cell 0 + Blot Module for Western blots was purchased from Novex. Fujifilm PPM-100A was from Eastman Kodak Co. All water was obtained from a Milli-Q system (Millipore Corp.), and all chemicals were of analytical grade unless otherwise stated.

**Purification of MBL** was performed by NatImmune A/S using a human embryonic kidney cell line HEK293, principally as described elsewhere (19). rhMBL was captured from the cell debris-deprived cell culture broth using a Glucosamine-Sepharose 4FF column (custom-designed resin from Amersham Biosciences) and subsequently eluted with TBS containing EDTA. The retained product was concentrated on a Source resin from Amersham Biosciences) and subsequently eluted with TBS broth using a Glucosamine-Sepharose 4FF column (custom-designed resin from Amersham Biosciences). The full process also included steps to remove DNA by nuclease treatment (Benzonase; Merck) and removal of virus (Planovia filters; Asahi Kasei Corp.).

**Western Blot**—The protein samples were separated using 3–8% SDS-PAGE from Novex. 20 ng of protein was loaded per lane, and electrophoresis was for 2 h at 125 V. The gel was blotted onto a prewetted 45-nm polyvinylidene difluoride membrane using the XCell II blot module and a blotting time of 135 min at 25 V. The membrane was blocked for 30 min in a Tris buffer (10 mM Tris, 150 mM NaCl, pH 8.0) supplemented with 0.1% Tween 20, and the primary antibody was applied and incubated for 3–12 h with gentle shaking. The membrane was washed and incubated with the secondary antibody for 90 min with gentle shaking. The protein bands were finally visualized using the SuperSignal West substrate (Pierce) and Fujifilms.

**Electron Microscopy**—Purified MBL was diluted and adsorbed to glow-discharged, carbon-coated copper grids at room temperature. Grids were then stained with 2% uranyl acetate and stained with two drops of freshly prepared 0.75% uranyl formate (Pältz & Bauer, Waterbury, CT). Specimens were inspected with a Philips Tecnai 12 electron microscope operated at 120 kV. Images were taken at a nominal magnification of 52,000 A using low dose procedures.

**Coating of Magnetic Beads**—The MPIO® Long Chain Alkylamine beads (CPG Inc.) were coated with 50 μg of porcine trypsin as previously described (20).

**Proteolytic Digestions**—Lyophilized rhMBL was reconstituted in 8 M urea and diluted to 1 mM in 50 mM NH₄HCO₃ containing 5% chymotrypsin (1-chloro-3-tosylamido-7-aminot-2-heptanone-treated; Worthington). The digestion mixture was incubated for 4 h at 37 °C, and the enzyme activity was quenched by freezing until separation by reversed phase HPLC. Lyophilized N-terminal of rhMBL was reconstituted in 8 M urea and diluted to 1 mM in 50 mM NH₄HCO₃ containing 50 μl of trypsin-coated magnetic beads, and the digestion was incubated for 10 min in the dark prior to the addition of trypsin-coated beads. The digestions were incubated for 24 h at 37 °C with shaking and quenched by removal of the magnetic beads and freezing until separation by reversed phase HPLC. Some lyophilized fractions of N-terminal rhMBL peptides were reconstituted in 5 μl of 10 mM NH₄HCO₃ and 1 μl of subtilisin (1 pmol/μl of 10 mM NH₄HCO₃ Calbiochem). The digestion encoding for 2 h at room temperature and quenched by micropurification and elution onto MALDI target.

**Reversed Phase HPLC**—All proteolytic digestions were separated using the tKta LC-900 system controlled by Unicorn software (version 3.21.02; Amersham Biosciences). All separations were performed on a Jupiter 5 μm C₄, 250 μm × 4.60-mm column (Phenomenex), using a 1%/min gradient from 5–50% B solvent followed by a 3%/min gradient from 50–80% B solvent. The A solvent used was 96% trifluoroacetic acid in water, and the B solvent was 90% acetonitrile, 0.05% trifluoroacetic acid in water. The subdigestions of the N-terminal part of rhMBL were separated using a Jupiter 5 μm C₄, 250 μm × 2.00-mm column (Phenomenex). Fractions were analyzed by MALDI MS or MALDI Q-TOF MS.

**Reduction of Disulfides**—Lyophilized samples were reconstituted in 10 mM DTT, incubated at 56 °C for 10 min, and analyzed by MALDI MS or MALDI Q-TOF MS.

**N-Terminal Sequencing**—N-terminal sequencing of peptides was performed on a Hewlett Packard G1005A protein sequencing system run according to the manufacturer’s recommendations. Approximately 100 pmol of sample was applied to the sequencer.

**MALDI MS of rhMBL Peptides**—MALDI MS analysis of peptides was performed on a Voyager-DE™ STR controlled by the Data Explorer software (version 3.4.0.0; both from PerSeptive Biosystems). The instrument was run with an acceleration voltage of 20 kV, grid voltage 65%, and delay time 100 ns. 100–300 shots were accumulated per spectrum. Approximately 1 pmol of sample was loaded in α-cyano-4-hydroxycinnamic acid (20 μg/μl in 70% acetonitrile, 0.1% trifluoroacetic acid) using the dried droplet method.

**MALDI Q-TOF MS**—MALDI Q-TOF MS of peptides was performed on a Q-Tof Ultima™ MALDI controlled by MassLynx software (version 3.5; both from Micromass Ltd.). Approximately 10 pmol of sample was concentrated on a reverse phase microcolumn (POROS® 50 R2; PerSeptive Biosystems) packed in a GearLoader tip (Eppendorf) as previously described (21). The sample was eluted onto the target with a 10 μg/μl solution of α-cyano-4-hydroxycinnamic acid in 70% acetonitrile, 0.1% trifluoroacetic acid.

**Data Analysis**—All data obtained on the Voyager-DE™ STR was processed using the m/z software (version 2001.08.14; Proteomics Inc.). External calibration of all spectra was performed using tryptic peptides of lactoglobulin (MH⁺ values of 837.48 and 2313.27). All MALDI Q-TOF data were processed using the MassLynx software (version 3.5; Micromass Ltd.). Peptides and modifications were identified using the GPMW software (version 5.03; Lighthouse Data) and the sequence of human MBL (Swiss-Prot number P11226), without the 20-aa signal peptide.

**Plasmid Construction and Site-directed Mutagenesis**—Human cDNA encoding wild-type MBL has previously been cloned (19). The cDNA encoding MBL was recloned into the pcDNA3.1 vector (Invitrogen). Briefly, the cDNA encoding for MBL is placed under the control of the cytomegalovirus promoter-enhancer.

To create the mutant variants of MBL, we mutated Cys³, Cys²⁵, and Cys²⁷ to Ser using a PCR-based method. We modified the QuickChange multisite-directed mutagenesis (Stratagene, La Jolla, CA) method to create all possible combinations of mutants. The mutagenesis reactions were carried out as described elsewhere. After mutagenesis, the relevant parts of the plasmid were sequenced to verify that the correct mutations were generated.

**Transient Transfection of HEK-293 Cells**—Transient transfection of the mutated MBL constructs into human embryonic kidney HEK293-F cells was performed using the FreeStyle™ 293 Expression System (Invitrogen). For transfection, 2 μg of each of the plasmids,complexed to the transfection reagent 293fect, was added to 5 × 106 cells in T-25 flasks with FreeStyle™ 293 Expression Medium (Invitrogen) supplemented with ascorbic acid. The transfected cells were incubated at 37 °C with a humid atmosphere of 8% CO₂. After 7 days, the supernatants were harvested and analyzed by Western blot.

**RESULTS**

**Oligomeric Distribution of MBL**—Electron microscope pictures of pMBL and rhMBL (Fig. 1) show that the structures present in the plasma-derived preparation of MBL are also present in the rhMBL preparation. A Western blot of rhMBL also shows the presence of the many different oligomer forms.

---

2 P. H. Jensen and D. Weilguny, manuscript in preparation.
Upon reduction all the forms of rhMBL are reduced to the polypeptide chain of MBL with the apparent mass of ~32 kDa (Fig. 2B).

Isolation of the N-terminal Part of rhMBL—In order to isolate the N-terminal part of rhMBL from the C-terminal part, several different digests (trypsin, endoproteinase Lys-C, collagenase) and chemical cleavage (cyanogen bromide) were set up (data not shown). However, the N-terminal peptides could not be recovered from any of these preparations. The N-terminal peptides were finally obtained from a chymotryptic digest separated by reversed phase HPLC (Fig. 3). All of the fractions were monitored by MALDI MS in order to locate the N-terminal region. The heterogeneous peak eluting at ~25% (v/v) acetonitrile (27–30 min in Fig. 3) contained Glu1-Leu46 of rhMBL. The mass spectrum inserted in Fig. 3 shows the reduced N-terminal peptide with an m/z value of 5316. The four peak clusters at m/z 5154, 4992, 4830, and 4668 correspond to the loss of 1–4 monosaccharide units respectively, originating from the glycosylated residues Lys36 and Lys39, probably due to in-source fragmentation. The heterogeneity of each peak cluster arises from the variation in hydroxylation of residues Pro21 and Pro27. The identity of the peptides was further confirmed using Edman degradation.

Identification of Disulfide-linked Peptides and the Bonds Involved—The disulfide-linked peptides Glu1-Leu46 were subdigested with trypsin, and the resulting peptide mix was separated by reversed phase HPLC. All peaks in Fig. 4 were monitored, reduced, and nonreduced by MALDI MS in order to locate disulfide-linked peptides. A tryptic digest of the N-terminal part of MBL yields two cysteine-containing peptides: Glu1–Lys10 and Thr11–Lys29. Several different disulfide-linked
species were identified in the reversed phase HPLC run as indicated in Fig. 4.

One fraction contained two peptides Glu–Lys10 interlinked by a disulfide bond (Fig. 4). Fig. 5A illustrates how the peak at m/z 2243 disappears upon reduction, yielding m/z 1123, which corresponds to Glu–Lys10. This identifies the presence of a Cys5–Cys5 bond in rhMBL as illustrated by the inset in Fig. 5A. The presence of a peak at m/z 1123 in the nonreduced spectrum is due to partial in-source fragmentation of the disulfide bond.

Several fractions contained two peptides Glu–Lys10 linked to one peptide Thr11–Lys29 (Fig. 4). This separation into multiple fractions reflects the heterogeneity in the hydroxylation of Pro21 and Pro27. As illustrated by Fig. 5C, the peak at m/z 3697 disappears upon reduction, yielding a peak at m/z 1849 corresponding to peptide Thr11–Lys29. This indicates the presence of a Cys5–Cys18 bond in rhMBL, as illustrated by the inset in Fig. 5C. The presence of a peak at m/z 1123 in the nonreduced spectrum is due to partial in-source fragmentation of the disulfide bond.

Several fractions contained two peptides Glu–Lys10 linked to one peptide Thr11–Lys29 (Fig. 4). The separation into multiple fractions reflects the heterogeneity in the hydroxylation of Pro21 and Pro27. As illustrated by Fig. 5B, the peak at m/z 4092 disappears upon reduction, yielding two peaks at m/z 1123 and 1849 corresponding to the peptides Glu–Lys10 and Thr11–Lys29. This identifies the presence of a Cys5–Cys12 and a Cys5–Cys18 bond in rhMBL as illustrated by the inset in Fig. 5B. Also present in the nonreduced spectrum in Fig. 5B is a peak at m/z 2972, which disappears upon reduction. This peak corresponds to one peptide Glu–Lys10 and one peptide Thr11–Lys29 interlinked by a disulfide bond. This species arises from the in-source fragmentation of a disulfide bond resulting in the loss of one peptide Glu–Lys10 (m/z 1123, also present in the spectrum) within the mass spectrometer.

Several fractions contained two interlinked peptides Thr11–Ala17 and two interlinked Cys18–Lys29 peptides, cleavage taking place at the Ala15–Cys16 bond. This identifies the presence of Cys12–Cys12 and Cys18–Cys18 bonds in rhMBL, as illustrated by the two building blocks inserted in the top of Fig. 6. Upon reduction, the peak at m/z 1196 (Cys18–Lys29) gets more intense, whereas the peak of Thr11–Ala17 (m/z 672) is below the cut-off m/z. This parameter is set to avoid suppression of the peptide signals, due to intense matrix signals in the low mass range.

Identification and Localization of a Free Cysteine—Fig. 4 also shows one fraction containing two peptides Glu–Lys10 linked to two peptides Thr11–Lys29. However, this species was not identified in the control run, where IAA was added prior to trypsin digestion. Here one peptide Glu–Lys10 linked to one peptide Thr11–Lys29 with one IAA was identified. Upon reduction of this species, two peaks at m/z 1123 and 1907 appeared, corresponding to Glu–Lys10 and Thr11–Lys29 + IAA (data not shown). The peak at m/z 1907 was further fragmented by MALDI Q-TOF MS, and from the resulting mass spectrum (Fig. 7) most of the peptide sequence could be deduced. This sequence reveals that Pro21 and Pro27 are both hydroxylated, Asn24 is deamidated (procedure-related artifact), and Cys18 is not modified by IAA, leaving only Cys12 to be alkylated, indicating the presence of a free Cys12 in rhMBL. This is illustrated by the inset in Fig. 7. Two such peptides can, if the cysteines are not protected otherwise, form a Cys12–Cys12 disulfide bond. This scrambling can give rise to the peptide Thr11–Lys29 and Cys12–Cys18 disulfide bond.

Also shown in Fig. 4 is the presence of three interlinked peptides Thr11–Lys29. Based on the intensity of the mass spectra, this tripeptide is present in less than 2% of the dipeptide. The tripeptide is no longer observed in the control run (i.e. after alkylation), indicating that it must have arisen by scrambling of a larger oligomer by the free cysteine (Cys12). The conformation of this species is likely to resemble that of two-cysteine collectins like CL-43, but due to the low amount present, this was not investigated.

Oligomerization of the Cys to Ser Mutants of rhMBL—The preliminary binding model proposed, based on the MALDI MS results shown so far, was challenged by the production of Cys to
Ser mutants of the N-terminal cysteines responsible for this binding. A series of PCR setups led to the production of all of the possible Cys to Ser mutants (i.e. the three single mutants, the three double mutants, and the triple mutant). A Western blot of their oligomerization patterns is illustrated in Fig. 8. The C5S and C12S single mutants form mainly polypeptide dimers and subunit monomers. C5S also forms a small amount of different oligomers. The C18S single mutant and all of the double mutants form polypeptide monomer but mainly polypeptide dimers. The triple mutant only exists as polypeptide monomers, as expected.

**Model for Oligomer Formation in rhMBL**—Since the MBL molecule is known to consist of 2–8 subunits plus intermediate forms (Fig. 2), a model will have to allow for the continuous assembly as well as the termination of disulfide bonding. The results from the MALDI MS experiments (Figs. 5–7) and the Cys to Ser mutants (Fig. 8) together gave rise to the oligomerization model illustrated in Fig. 9. This scheme proposes the formation of two isomers of a continuous subunit as well as a terminating subunit.

**DISCUSSION**

The present study was performed using recombinant human mannann-binding lectin (rMBL) as opposed to plasma-derived mannann-binding lectin (pMBL). This approach presents several advantages. First, rMBL produced from one clone expressing a single type of MBL is more homogenous than pMBL purified from a pool of plasma, where several isoforms are known to exist (15–17). Second, the quantities necessary for this kind of study are readily available, unlike performing a similar study on pMBL. Purification of such a large amount of homozygous wild type MBL would require large amounts of genotyped A/A plasma, and interpretation of the results could be further complicated by the presence of partially degraded pMBL as a result of the natural turnover in the human body. In addition, all available evidence indicates that no major differences exist, with regard to post-translational modification and processing, between the native and the recombinant protein. The rMBL is produced in human cells, from a construct containing the wild type human MBL gene. A previous study of the similarity between pMBL and rMBL found that HEK293 rMBL is both structurally and functionally similar to pMBL (19). This study analyzed both the degree of oligomerization and the complement activating properties of rMBL. Electron microscopy pictures comparing the structures of pMBL and rMBL further support the high similarity (Fig. 1). Furthermore, all post-translational modifications located in rMBL have also

---

**Fig. 5. Identification of disulfide bonds in rhMBL.** A, the top panel shows the MALDI MS spectrum of the nonreduced sample (+DTT), and the lower panel shows the corresponding reduced sample (-DTT). Upon reduction, m/z 2243 disappears, giving rise to m/z 1123, which corresponds to Glu1–Lys10 of rhMBL. m/z 2243 corresponds to two interlinked Glu1–Lys10 peptides, shown as a building block inserted in the lower right corner. B, upon reduction, m/z 4092 and 2972 disappear, giving rise to m/z 1849 and 1849, which corresponds to Glu1–Lys10 and Thr11–Lys29 of rhMBL. m/z 4092 corresponds to two Glu1–Lys10 peptides and one Thr11–Lys29 peptide. 2972 arises from the loss of Glu1–Lys10 in the mass spectrometer. The corresponding building block is inserted in the lower right corner. C, upon reduction, m/z 3697 disappears, giving rise to m/z 1849, which corresponds to Thr11–Lys29 of rhMBL. m/z 3697 corresponds to two interlinked Thr11–Lys29 peptides, shown as a building block inserted in the lower right corner. Since the peptide Thr11–Lys29 contains two cysteines (Cys12 and Cys18), the disulfide bonds cannot be deduced from these spectra.

**Fig. 6. Identification of a Cys12–Cys18 and a Cys18–Cys29 disulfide bond in rhMBL.** The fraction containing the m/z 3697 was subdigested with subtilisin as described under “Experimental Procedures.” The top panel shows the MALDI MS spectrum of the original sample, the middle panel shows the nonreduced, subtilisin-digested sample, and the lower panel shows the corresponding reduced sample. The digestion gives rise to m/z 1345 and 2390, which disappear upon reduction. m/z 1345 corresponds to two interlinked Thr11–Ala17 peptides, and m/z 2390 corresponds to two interlinked Cys18–Lys29 peptides, shown as building blocks inserted in the lower right corner. Upon reduction, m/z 1196 (Cys18–Lys29) is seen, whereas m/z 672 (Thr11–Ala17) is lost in the matrix noise.
been identified in pMBL (23), and the complement activating capabilities of the two preparations are similar (24).

The MBL subunit consists of three interlinked polypeptide chains (the subunit monomer) that further assemble into oligomeric forms. The collagen-like region, in combination with the \( \alpha \)-helical coiled-coil neck region, is responsible for the association of the polypeptide chains into the subunit monomer (10, 11). The subunit monomer is further stabilized by intrasubunit disulfide bonds in the N-terminal region (12). Inter-subunit disulfide bonds in the N-terminal region have been shown to be responsible for the association of the monomer subunits into oligomeric forms (14). The function of MBL in the immune system is tightly linked to the level of oligomerization. The resulting small oligomer structures are believed to be responsible for opsonizing pathogens, whereas the larger forms activate complement (6–8).

Several species, including mice and rats, produce two types of MBL, one type containing two N-terminal cysteines, which does not oligomerize, and one type containing three N-terminal cysteines, which forms higher oligomers like the human MBL species does (22, 25, 26). The N-terminal disulfide-bonding pattern of the rat MBL containing two cysteines (MBL-C) was solved by Wallis and Drickamer (22). This model involves asymmetrical bonds, indicating a large flexibility of the polypeptide chains in the subunit. This pattern is identical to that determined for CL-43 (27). The N-terminal disulfide binding pattern has not previously been elucidated for any of the MBLs containing three cysteines, apart from the observation that the first of the three N-terminal cysteines seems to be responsible for the oligomerization, whereas the two other cysteines form intrasubunit bonds (25, 26).

Due to the large heterogeneity of the MBL polypeptide chain, a crucial aspect of the identification of the disulfide-bonding pattern was the isolation of the “intact” N-terminal region. Collagenase has been used to isolate intact and functional CRDs of MBL (28); however, we have been unable to isolate the

---

**Fig. 7. Localization of a free cysteine in rhMBL.** Free cysteines were localized by the addition of IAA prior to the digestion of Glu1–Leu46. One modified peptide with the \( m/z \) of 1907 was identified by MALDI MS. This corresponds to Thr11–Lys29 + IAA. The peptide was sequenced by MALDI Q-TOF MS, yielding the sequence inserted at the top. It shows that Pro21 and Pro37 are hydroxylated, Asn38 is deamidated, and Cys46 is unmodified. This indicates that the modified residue is Cys12, as illustrated by the building block inserted below the sequence. *, the \( Y \) ion series of the sequence, which is read from right to left.

**Fig. 8. Western blot of the Cys to Ser mutants of rhMBL.** The oligomerization pattern of the seven Cys to Ser mutants was visualized by nonreducing SDS-PAGE and Western blot. B, the overexposed equivalent to A. The blot shows that CSS is the only mutant capable of oligomerizing, however only to a small extent. The C12S can form the subunit monomer (SM), whereas the C18S and the double mutants merely form the polypeptide dimer (PD). The triple mutant only forms polypeptide monomers (PM) as expected. The molecular marker mass values are given in kDa. WT, wild type.
The N-terminal region using this approach. Most other specific enzymes (trypsin, endoproteinase Lys-C, and Asp-N) cleave between the first two cysteines, yielding smaller peptides, and except for a tryptic dimer of the first Cys residue, no cross-linked peptides from the N-terminal region were ever recovered. Only digestion with chymotrypsin resulted in recovery of the various intact forms of the cross-linked N-terminal region (Fig. 3). Neither MALDI MS nor SDS-PAGE yielded any information about the number of interacting polypeptides. This is most likely due to the high heterogeneity of the samples, which is clearly seen in the HPLC chromatogram. Subdigestion of the chymotryptic peptides with trypsin and subtilisin resulted in the identification of several specific disulfide bonds. The cleavages abolished information on inter- versus intrasubunit disulfide bonds, but the peptides identified gave rise to a set of building blocks that form the oligomer structures of rhMBL (Figs. 5–7).

The Cys to Ser mutants and their disulfide bonding patterns (Fig. 8) offer support to the results obtained by MALDI MS and show that the building blocks represent both intra- and inter-subunit disulfide bonds. The lack of one, two, or three cysteines simplifies the interpretation of the very complex bonding pattern in question. The single mutants indicate the role of each of the three cysteines in the complex oligomerization process of rhMBL. The C5S is the only mutant that oligomerizes, whereas C12S forms subunit monomers and C18S only forms polypeptide dimers. The double mutants offer support for the existence of the Cys$^5$–Cys$^9$, Cys$^{12}$–Cys$^{12}$, and Cys$^{18}$–Cys$^{18}$ bonds identified by MALDI MS. The triple mutant stresses the importance of the N-terminal cysteines in the folding of active rhMBL, since this mutant only exists as polypeptide monomers.

The results reported here can be used to explain the assembly of the different MBL oligomer structures. From the building blocks (Figs. 5–7) and the bonding patterns of the Cys to Ser mutants (Fig. 8), a scheme for the assembly of the ultrastructures of rhMBL is proposed (Fig. 9). The scheme leads to the formation of one of three subunit structures, which decide the further “fate” of this particular rhMBL. The model indicates that the preferred binding pattern is symmetrical and repetitive but leaves room for the termination of the oligomerization at any point and is thus able to account for all of the different oligomer forms of MBL.

The dynamics of the proposed oligomer assembly are similar to the ones previously described by Weis and colleagues (1). We speculate that as the collagen-like region zips from the C terminus (11) and the N-terminal cysteines come into contact, the first Cys$^{18}$–Cys$^{18}$ bond forms. Afterward, the Cys$^{12}$–Cys$^{12}$ bond forms, preferably involving the third polypeptide, after which a Cys$^{5}$–Cys$^{5}$ bond is formed. This leaves three “free” cysteines to interact with neighboring subunits.

There are two isomers of this continuous subunit, varying only in the polypeptides participating in the Cys$^{5}$–Cys$^{5}$ bond (Fig. 9). Either of the two isomers satisfies the building blocks illustrated in Figs. 5A and 6. However, since they are at the peptide level, these building blocks yield no information of the number of intact polypeptides involved. This information comes from the Cys to Ser single mutants. Both isomers allow C5S to form subunit monomer and leave the potential to oligomerize, which this mutant does. Isomer 1 allows C12S to form polypeptide dimer, whereas isomer 2 allows it to form subunit monomer. It also leaves the potential to oligomerize, but since this is not the case (Fig. 8), the binding pattern identified in CL-43 and rat MBL-C is likely to be applicable here. For C18S, isomer 2 accounts for the formation of polypep-
tide dimers, whereas isomer 1 indicates that it should be able to form subunit monomers. The absence of this structure in Fig. 8 may be due to C18 playing an important role in the initiation of the correct bonding.

A terminating subunit (Fig. 9) may occasionally form if the Cys12–Cys12 bond forms between the two polypeptides linked by the Cys8–Cys18 bond. This leads to the third polypeptide end flipping down along the side of the structure forming the Cys5–Cys18, thereby terminating the oligomerization process. The Cys2–Cys12 bond of this terminating subunit arises as the folding of one chain makes Cys12 available for binding to the Cys5 of a neighboring polypeptide. This terminating subunit satisfies the building block in Fig. 5B and also accounts for the polypeptide dimer formation seen in the C5S mutant in Fig. 8. Based on the absorption of the different building blocks in Figs. 4 and the low amounts in gels (Fig. 2), we speculate that this termination pattern does not happen very often. Alternatively, termination of oligomerization can presumably happen by circularization of a number of continuous subunits. Considering all of the possible different oligomer forms, this process is proposed to be somewhat random in time but may be dependent on the concentration during synthesis. Sometimes polypeptide monomers or dimers (disulfide-linked), not wound in a collagen coil, could join the disulfide bonding and account for intermediate oligomers. These intermediate forms are not as common as the full subunit forms (Fig. 2). Likewise, it is not known whether both isomer 1 and isomer 2 are present in the final MBL structure.

Fig. 7 shows the identification of the free cysteine. Although the presence of a free cysteine in an uneven oligomer is a prerequisite, previous structural studies failed to locate free cysteines (26, 29). One explanation is the small amount of potential free cysteine compared with the total amount of cysteine. The free cysteines would account for less than 1% of the total Cys residue content, since in the proposed model, 1 of 105 Cys residues in a pentamer is expected not to participate in MBL bonding. In a mixture with even numbered oligomers, this ratio is expected to be even lower. Another explanation could be that the free cysteine in vivo is likely to be modified by free cysteine or glutathione. The two continuous subunits in the model in Fig. 9 each have three “free” cysteines. At least two cysteines are likely to be participating in the oligomerization with neighboring subunits. In oligomers with an even number of subunits, all three cysteines most likely participate in disulfide bonding. However, in the case of an uneven numbered oligomer, one cysteine will be unable to participate, and both isomers allow for Cys12 to be free in accordance with the building block in Fig. 7. An explanation for why only Cys12 appears to end up as a free cysteine must await the purification of pure oligomeric forms.

The ultrastructure of MBL is often compared with C1q, the first component of the classical pathway of complement. The two molecules are similar both in function and structure, but they also have their differences. C1q is composed of six heterotrimers, linked two and two by disulfide bonds in the N-terminal region. A nonreducing SDS-PAGE of C1q shows only two bands, corresponding to the A-B polypeptide dimer and the C-C polypeptide dimer (30). A nonreducing SDS-PAGE of MBL shows a large number of different oligomer forms, indicating that the MBL polypeptides and subunits are covalently interlinked (26, 31, 32), in contrast to the collagen-like structure and noncovalent interactions holding together the C1q hexamer. This is supported by the fact that the Cys to Ser triple mutant only exists as single polypeptide chain and abolishes the possibility of superimposing the N-terminal bonding pattern of C1q upon MBL.

MBL is structurally closely related to the other collectins. At the polypeptide level, all of the collectins have a similar domain organization. All of the collectins form homotrimeric, but their oligomer structures divide them into several classes. CL-43 (33) and MBL-C (22) exist only as monomers, whereas surfactant protein D (34) and conglutinin (35) associate into cruciformed tetramers. Surfactant protein A (36) forms a hexamer bouquet-like structure similar to C1q and MBL. Not much is known about the ultrastructure of the two most recently discovered collectins: liver collectin 1 (37) and CL-46 (38). It is believed that the N-terminal cysteines and their bonding patterns are responsible for the variation in the oligomer structures of the collectins. Indeed, CL-43 and the nonoligomerizing form of rat MBL (MBL-C) show identical bonding patterns (22, 27). The bonding patterns for the other collectins have not been solved. The known pattern cannot be directly superimposed on human MBL, since this protein and the rat oligomer form (MBL-A) are the only collectins containing three N-terminal cysteines. The C5S and C12S single mutants show an oligomer distribution that could indicate a binding pattern similar to that described for CL-43 and rat MBL-C. The fact that the C5S mutant behaves in the same way as the C12S mutant is in contrast to previous studies stating that the first cysteine is responsible for intersubunit bonds, whereas the next two form intrasubunit bonds (25, 26). The fact that the C18S does not form the subunit monomer could indicate either that there is steric hindrance preventing the formation of the expected Cys5–Cys12 bond or that isomer 2 of the continuous subunit is more common (Fig. 9). This study shows that the number as well as the position of cysteines determines the oligomerization pattern of MBL.

The three well characterized mutations in the collagen-like region of human MBL (MBL-B, -C, and -D) are known to decrease or completely abolish the complement-activating activity of the protein (7, 8, 39, 40). The mechanism for this is believed to be altered or impaired oligomerization, and determination of the N-terminal disulfide-bonding pattern of these variants could answer this question. Indeed, MBL-D is characterized by the introduction of an extra cysteine in the collagen-like domain. This extra cysteine results in an even number and thereby abolishes the need for oligomerization in order to form a stable molecule.

Acknowledgments—We thank T. Walz and B. Petre for assistance with electron microscopy.

REFERENCES

1. Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992) Nature 360, 127–134
2. Turner, M. W. (1996) Immunol. Today 17, 532–540
3. Matsushita, M., and Fujita, T. (1992) J. Exp. Med. 176, 1497–1502
4. Dahl, M. R., Thiel, S., Matsushita, M., Fujita, T., Willis, A. C., Christensen, T., Vorup-Jensen, T., and Jensenius, J. C. (2001) Immunity 15, 127–135
5. Thiel, S., Vorup-Jensen, T., Stover, C. M., Schwaebel, W., Laursen, S. B., Poulsen, K., Willis, A. C., Eggeler, P., Hansen, S., Holmskov, U., Reid, K. B., and Jensenius, J. C. (1997) Nature 386, 506–510
6. Holmskov, U., Laursen, S. B., Malhotra, R., Wiedemann, H., Timpl, R., Stuart, G. R., Torneoe, I., Madsen, P. S., Reid, K. B., and Jensenius, J. C. (1995) Biochem. J. 305, 889–896
7. Kurata, H., Cheng, H. M., Kosztuzumi, Y., Yokota, Y., and Kawasaki, T. (1993) Biochem. Biophys. Res. Commun. 191, 1204–1210
8. Sugai, M., Gillies, S. D., Foley, S., Saxtry, K., Schweinle, J. E., Silverman, V. J., and Eozkowiz, B. A. (1992) Nat. Genet. 2, 50–55
9. Drickamer, K., and Taylor, M. E. (1993) Annu. Rev. Cell Biol. 9, 237–264
10. Ma, Y., Shida, H., and Kawasaki, T. (1997) J. Biochem. (Tokyo) 122, 810–818
11. Chida, R., Feizi, T., Yamada, T., Drickamer, K., and Quisenberry, S. (1990) J. Biol. Chem. 265, 20770–20777
12. Hansen, S., and Holmskov, U. (1998) Immunochemicals 199, 165–189
13. Ikeda, R., Sannoh, T., Kawasaki, T., and Yamashina, I. (1987) J. Biol. Chem. 262, 7451–7454
14. Hoppe, H. J., and Reid, K. B. (1994) Protein Sci. 3, 1143–1158
15. Garred, P., Thiel, S., Madsen, H. O., Ryder, L. P., Jensenius, J. C., and Svegaard, A. (1995) Clin. Exp. Immunol. 96, 517–521
The Oligomer Structure of rhMBL

11051

S., and Svejgaard, A. (1994) Immunogenetics 40, 37–44
18. Madsen, H. O., Garred, P., Thiel, S., Kurtzhals, J. A., Lamm, L. U., Ryder, L. P., and Svejgaard, A. (1995) J. Immunol. 155, 3013–3020
19. Vorup-Jensen, T., Sørensen, E. S., Jensen, U. B., Schwaeble, W., Kawasaki, T., Ma, Y., Uemura, K., Wakamiya, N., Suzuki, Y., Jensen, T. G., Takahashi, K., Ezekowitz, A. B., Thiel, S., and Jensenius, J. C. (2001) Int. Immunopharmacol. 1, 677–687
20. Krogh, T. N., Berg, T., and Højrup, P. (1999) Anal. Biochem. 274, 153–162
21. Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., and Roepstorff, P. (1999) J. Mass Spectrom. 34, 105–116
22. Wallis, R., and Drickamer, K. (1997) Biochem. J. 325, 391–400
23. Nielsen, P. H. (2001) Primary Structure Analysis of Human MBL: Disulfide Bonds and Other Modifications. Master’s thesis, University of Southern Denmark, Odense, Denmark
24. Jensenius, J. C., Jensen, P. H., McGuire, K., Larsen, J. L., and Thiel, S. (2003) Biochem. Soc. Trans. 31, 763–767
25. Sastry, R., Wang, J. S., Brown, D. C., Ezekowitz, R. A., Tauber, A. I., and Sastry, K. N. (1995) Mamm. Genome 6, 103–110
26. Wallis, R., and Drickamer, K. (1999) J. Biol. Chem. 274, 3580–3589
27. Rothman, A. B., Mortensen, H. D., Holmskov, U., and Højrup P. (1997) Eur. J. Biochem. 243, 630–635
28. Drickamer, K., Dordal, M. S., and Reynolds, L. (1986) J. Biol. Chem. 261, 6878–6887
29. Jensen, U. G. (1996) Purification and Characterization of the Human Lectin Mannan-binding Protein. Master’s thesis, University of Odense, Denmark
30. Reid, K. B., and Porter, R. R. (1976) Biochem. J. 155, 19–23
31. Liu, H., Jensen, L., Hansen, S., Petersen, S. V., Takahashi, K., Ezekowitz, A. B., Hansen, P. D., Jensenius, J. C., and Thiel, S. (2001) Scand. J. Immunol. 53, 489–497
32. Lu, J. H., Thiel, S., Wiedemann, H., Timpl, R., and Reid, K. B. (1990) J. Immunol. 144, 2287–2294
33. Holmskov, U., Teisner, B., Wallis, A. C., Reid, K. B., and Jensenius, J. C. (1993) J. Biol. Chem. 268, 10120–10125
34. Persson, A., Chang, D., Rust, K., Moxley, M., Longmore, W., and Crouch, E. (1989) Biochemistry 28, 6361–6367
35. Davis, A. E., III, and Lachmann, P. J. (1984) Biochemistry 23, 2139–2144
36. White, R. T., Damm, D., Miller, J., Spratt, K., Schilling, J., Hawgood, S., Benson, B., and Cordell, B. (1985) Nature 316, 361–363
37. Ohnani, K., Suzuki, Y., Eda, S., Kawai, T., Kase, T., Yamazaki, H., Shimada, T., Keshi, H., Sakai, Y., Fukuhk, A., Sakamoto, T., and Wakamiya, N. (1999) J. Biol. Chem. 274, 13681–13689
38. Hansen, S., Holm, D., Moeller, V., Vêved, L., Bendixen, C., Reid, K. B., Skjoedt, K., and Holmøvak, U. (2002) J. Immunol. 169, 5726–5734
39. Wallis, R., and Cheng, J. Y. (1999) J. Immunol. 163, 4953–4959
40. Larsen, F., Madsen, H. O., Sim, R., Koch, C., and Garred, P. (2004) J. Biol. Chem. 279, 21302–21311
Characterization of the Oligomer Structure of Recombinant Human Mannan-binding Lectin

Pia H. Jensen, Dietmar Weilguny, Finn Matthiesen, Kirsten A. McGuire, Lei Shi and Peter Højrup

J. Biol. Chem. 2005, 280:11043-11051.
doi: 10.1074/jbc.M412472200 originally published online January 14, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M412472200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 36 references, 13 of which can be accessed free at http://www.jbc.org/content/280/12/11043.full.html#ref-list-1