Asymmetry adjacent to the collagen-like domain in rat liver mannose-binding protein

Russell WALLIS and Kurt DRICKAMER*

Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Rat liver mannose-binding protein (MBP-C) is the smallest known member of the collectin family of animal lectins, many of which are involved in defence against microbial pathogens. It consists of an N-terminal collagen-like domain linked to C-terminal carbohydrate-recognition domains. MBP-C, over-produced in Chinese-hamster ovary cells, is post-translationally modified and processed in a manner similar to the native lectin. Analytical ultracentrifugation experiments indicate that MBP-C is trimeric, with a weight-averaged molecular mass of approx. 77 kDa. The rate of sedimentation of MBP-C and its mobility on gel filtration suggest a highly elongated molecule. Anomalous behaviour on gel filtration due to this extended conformation may explain previous suggestions that MBP-C forms a higher oligomer. The polypeptide chains of the MBP-C trimer are linked by disulphide bonds between two cysteine residues at the N-terminal junction of the collagen-like domain. Analysis of an N-terminal tryptic fragment reveals that the disulphide bonding in MBP-C is heterogeneous and asymmetrical. These results indicate that assembly of MBP-C oligomers probably proceeds in a C- to N-terminal direction: trimerization at the C-terminus is followed by assembly of the collagenous domain and finally formation of N-terminal disulphide bonds. The relatively simple organization of MBP-C provides a template for understanding larger, more complex collectins.

INTRODUCTION

The collectins are a family of Ca^{2+}-dependent (C-type) animal lectins that mediate host defence against pathogens. The best-characterized collectins are serum mannose-binding proteins (MBPs) which bind carbohydrates on the surfaces of pathogenic micro-organisms, induce opsonization and activate complement in an antibody-independent manner [1–3]. Each collectin polypeptide consists of an N-terminal cysteine-rich domain, a collagen-like region, a neck and a C-terminal C-type carbohydrate-recognition domain (CRD) [4]. These polypeptide chains associate to form homotrimers which in turn are assembled into higher order oligomers.

Defined by their oligomeric structures, most collectins fall into one of two groups [4]. Those with shorter collagenous domains of 18–24 Gly-Xaa-Yaa repeats, including mouse, rat and human serum MBPs and pulmonary surfactant Protein A, contain up to six trimeric units. These oligomers form bouquet-like structures that appear similar to complement component C1q. In contrast, conglutinin and pulmonary surfactant protein D (SP-D) contain up to 59 Gly-Xaa-Yaa repeats and are cruciform. Rat liver mannose-binding protein (MBP-C) is most similar to members of the former group based on its amino acid sequence as it shares 56 % identity with rat serum mannose-binding protein (MBP-A) [5]. Gel-filtration data have been interpreted to indicate that MBP-C consists of two trimeric units [6].

The assembly of collectin oligomers is defined by both the primary structure and covalent modifications of the constituent polypeptides. For example, the collagen-like domain contains 4-hydroxyproline as well as 5-hydroxylysine residues, which are further modified by glycosylation to form glucosylgalactosyl-5-hydroxylysine [5,7]. As in vertebrate collagen, the hydroxyproline residues in the collectins are believed to stabilize the triple helix, which appears as an extended rod in electron micrographs [8]. A Gly-Gln-Gly interruption in the collagenous region of the MBPs is thought to form a kink in the triple helix which allows the trimers to bend away from each other in the 'bouquet-like' structure of MBP-A [5]. Analysis of collectins under reducing and non-reducing conditions reveals that the polypeptide chains and in some cases the trimeric building blocks are linked by disulphide bonds formed between the N-terminal cysteine-rich domains [6,8–11].

The organization of the CRD and neck fragments of MBP-A derived from crystallography provides an important starting point in the overall analysis of collectin molecular architecture. The CRDs are organized in a cluster stabilized by a coiled-coil of α-helices formed by the neck region [12]. The relationship between this portion of the molecule and the collagen and N-terminal domains remains to be established. In the present study, MBP-C is shown to be a simple trimer of intact collectin polypeptides, thus making it a good model for the complete trimeric building block of the collectins. To facilitate analysis of MBP-C structural organization, an expression system has been developed to generate material that faithfully reflects the properties of natural MBP-C. Analysis of N-terminal tryptic fragments, spanning the cysteine-rich domain, reveals that the three polypeptide chains must be organized in a heterogeneous and asymmetrical arrangement.

EXPERIMENTAL

Materials

Restriction enzymes and buffers were purchased from New England Biolabs. All tissue-culture medium was from Life Technologies. Phenylthiohydantoin (PTH)-amino acids and protein molecular-mass markers for gel filtration, SDS/PAGE and matrix-assisted laser desorption ionization mass spectrometry

Abbreviations used: MBP, mannose-binding protein; MBP-A, serum mannose-binding protein; MBP-C, liver mannose-binding protein; SP-D, pulmonary surfactant protein D; CRD, carbohydrate-recognition domain; MALDI-MS, matrix-assisted laser desorption ionization mass spectrometry; DTT, dithiothreitol; PTH, phenylthiohydantoin.

* To whom correspondence should be addressed.

(MALDI-MS) were purchased from Sigma. Tosylphenylalanylchloromethane ('TPCK')-treated trypsin and subtilisin were obtained from Worthington Biochemicals and Boehringer Mannheim respectively. Reagents for amino acid sequencing were purchased from Beckman Instruments. Iodo[2-^{3}H]acetic acid, reagents for silver staining and Amplify fluorography reagent were obtained from Amersham.

Analytical methods

Amino acid sequencing was carried out on a Beckman LF3000 protein sequencer. L-Cysteine hydrochloride, L-cystine and bovine insulin were subjected to Edman degradation to establish the elution position of di(PTH)cystine. Amino acid analysis was performed by the method of Heinrikson and Meredith [13]. SDS/PAGE was performed by the method of Laemmli [14]. Standard molecular-biology techniques were carried out as described in Sambrook et al. [15].

Preparation of MBP-C from liver

MBP-C was purified on a column of mannose–Sepharose from a detergent extract of rat liver acetone powder, based on procedures previously described [5,6]. The extract was passed through a galactose–Sepharose column to remove residual asialoglycoprotein receptor and the resulting protein was concentrated on a 1 ml mannose–Sepharose column. MBP-A was removed by gel-filtration chromatography using a TSK G3000SW column equilibrated in 100 mM potassium phosphate buffer, pH 7.0, containing 2.5 mM EDTA. Aliquots of 100 μl were loaded on to the column. Fractions containing MBP-C were identified at each step by SDS/PAGE.

Overproduction of MBP-C

The cDNA encoding MBP-C was modified to introduce a *Sal*I restriction site adjacent to the *Bst*YI site near the 5′-end (base 114) and an *Xho*I site at the *Bst*EII site in the 3′-untranslated region (base 1002). This fragment was ligated into the unique *Sal*I restriction site of the vector pED [16]. Since plasmid pED contains the dihydrofolate reductase gene downstream of this *Sal*I site, translation from the adenovirus major late promoter results in the production of a dicistronic mRNA. The resulting construct (pED-C) was used to transfect adherent Chinese-hamster ovary cells DXB11, deficient in dihydrofolate reductase [17], using the calcium phosphate method [18]. Cells containing the construct were selected by growth in minimum essential medium α lacking nucleosides and supplemented with 10% (v/v) dialysed foetal-calf serum. Production of MBP-C was increased over several weeks by passaging the cells into medium containing increasing concentrations of methotrexate (0.02, 0.1 and 0.5 μM). Cells were grown to confluence in 225 cm^2 tissue-culture flasks containing 50 ml of medium with 0.5 μM methotrexate. Medium was harvested every second day, and loaded on to a mannose–Sepharose column (1 ml) equilibrated in loading buffer [50 mM Tris/HCl (pH 7.5)/125 mM NaCl/25 mM $CaCl_2$] following the addition of 2.5 ml of 1 M Tris/HCl, pH 7.4, and 1.25 ml of 1 M $CaCl_2$. The column was washed with 10 ml of high-salt buffer (loading buffer containing 1.25 M NaCl) and 10 ml of loading buffer, before elution with 50 mM Tris/HCl (pH 7.5)/150 mM NaCl/2.5 mM EDTA. Where necessary, MBP-C in 20 mM Mes, pH 6.0, containing 10 mM EDTA, was further purified on a HR 5/5 MonoS cation-exchange column (Pharmacia) equilibrated in the same buffer. The protein was eluted with a 500 mM NaCl gradient over 45 min and fractions containing MBP-C were identified by SDS/PAGE.

As discussed below, MBP-C purified from the culture medium of confluent cells was found to resemble native lectin by a variety of biochemical and biophysical techniques. However, the apparent molecular mass of protein isolated from the medium of non-confluent cells was decreased. The difference in the preparations was not due to proteolysis at the N-terminus of MBP-C and may reflect differences in post-translational modification. For this reason, characterization of overproduced MBP-C was always carried out using protein purified from the culture medium of confluent cells.

Gel filtration

Analytical gel-filtration chromatography was carried out on a TSK G3000SW column equilibrated in 100 mM potassium phosphate buffer, pH 7.0, containing 2.5 mM EDTA. Bovine thyroglobulin, horse spleen apoferritin, sweet-potato β-amylase, yeast alcohol dehydrogenase, BSA and carbonic anhydrase were used as molecular-mass standards. The Stokes radius of MBP-C was calculated from its elution position using the method of Laurent and Killander [19]. Values for the protein standards were calculated from their diffusion coefficients [20].

Analytical ultracentrifugation

All experiments were carried out in a Beckman Optima XL-A analytical ultracentrifuge equipped with absorbance optics using an An60Ti rotor. Before setting up the cells, samples were dialysed overnight against the appropriate buffer [50 mM Tris/HCl (pH 7.5) containing 150 mM NaCl and 2.5 mM $CaCl_2$, unless otherwise stated]. For equilibrium experiments, samples (110 μl) were placed in the sample cells of an Epon charcoal-filled six-channel centrepiece. The reference cells were loaded with aliquots of the dialysate (125 μl). Experiments were performed at 8000–12000 rev./min and 20 °C.

Equilibrium data were collected at 280 nm in step scan mode using a separation of 0.001 cm. Five readings were averaged for each scan and a baseline scan was taken at 360 nm to correct for optical imperfections. Readings were taken at 4 h intervals until no difference could be detected between consecutive scans. The equilibrium distributions from three different loading concentrations and up to three different rotor speeds were analysed simultaneously using the Nonlin curve fit algorithm supplied with the centrifuge [21]. Data were also analysed by calculation of the apparent weight-averaged molecular mass at concentrations throughout the sample cells.

Sedimentation-velocity experiments were carried out at 42000 rev./min using an Epon aluminium-filled centrepiece. Sample (400 μl) and dialysate (425 μl) were loaded in the sample and reference cells respectively. Scans were collected at 5 min intervals in the step scan mode. The data were analysed by the second-moment method using the software supplied by Beckman Instruments. The partial specific volume of MBP-C (0.7177 ml/g) was calculated from the amino acid and carbohydrate composition using the method of Cohn and Edsall [22].

Light scattering

The diffusion coefficient of MBP-C (0.1–0.5 mg/ml) in 50 mM Tris/HCl, pH 7.5, containing 500 mM NaCl and 2 μM EDTA at 24.5 °C, was estimated by light scattering using a BioTage model 801 molecular-size detector (version 2.2).

Hydroxylamine cleavage

Aliquots of MBP-C (100 μl at 0.2–0.5 mg/ml) in 120 mM potassium carbonate, pH 9.0, were incubated with 2 M hydroxylamine at 45 °C for 4 h [23]. The hydroxylamine was removed by dialysis overnight against 50 mM Tris/HCl, pH 7.5, containing 150 mM NaCl. To precipitate the protein, trichloroacetic acid was added to a final concentration of 10 % (w/v) and the samples incubated on ice for 10 min. Following centrifugation at 18000 ***g*** in an Eppendorf centrifuge the protein pellets were rinsed with a 1:1 (v/v) mixture of ethanol and ether, freeze-dried and the cleavage products were separated by SDS/PAGE. Fragments were blotted on to a PVDF membrane for sequence analysis [24].

Determination of free thiol groups

Protein (1–5 μM) was incubated with a minimum 1000-fold excess of iodo[2-^{3}H]acetic acid (5 mM; 178 mCi/mmol) in 50 mM Tris/HCl, pH 7.5, containing 150 mM NaCl and 25 mM $CaCl_2$, in both the presence and absence of guanidine hydrochloride (4 M). After incubation for 1 h at 37 °C, the guanidine was removed by dialysis overnight against reaction buffer without iodoacetamide and the sample was precipitated with trichloroacetic acid (see above). The oligomeric species were separated by SDS/PAGE under non-reducing conditions. After staining and destaining, the gel was soaked in Amplify, dried and exposed to Kodak X-Omat AR-5 film for up to 1 week at −70 °C.

Trypsin digestion

MBP-C (1.5 ml at 1–2 mg/ml) in digestion buffer [50 mM Mops (pH 6.5) containing 25 mM $CaCl_2$] was incubated with 10 % (w/w) trypsin at 37 °C. After 3 h the reaction was stopped by the addition of a 1.2-fold excess of soya bean trypsin inhibitor and loaded on to a mannose–Sepharose affinity column (0.5 ml; pre-equilibrated in digestion buffer) to remove the CRDs which are relatively resistant to digestion. The column was washed with 3 column vol. of buffer and the eluate freeze-dried. Peptides were resuspended in 200 μl of 20 % (v/v) acetic acid and separated by reversed-phase chromatography.

Reversed-phase HPLC

Peptides were purified on a C3 reversed-phase column (5 μm; 4.6 mm × 7.5 cm; Beckman) equilibrated in 0.1 % (v/v) trifluoroacetic acid in water using a 0–50 % gradient of 0.1 % (v/v) trifluoroacetic acid in acetonitrile over 50 min at a flow rate of 1 ml/min. The absorbance was monitored at both 250 and 210 nm and 1 ml fractions were collected. When necessary, further purification was achieved by re-applying fractions to the column and eluting using a shallow gradient [0.5 % (v/v) acetonitrile/min].

MALDI-MS

Samples were analysed on a FinniganMAT Lasermat mass spectrometer. Peptides isolated by reversed-phase chromatography were freeze-dried and resuspended in water to a final concentration of 1–10 pmol/μl. Samples (0.5 μl) were spotted on to a target and mixed with 0.5 μl of α-cyano-4-hydroxycinnamic acid (7 mg/ml in 70 % acetonitrile containing 0.1 % trifluoroacetic acid). Reduced fractions were prepared by adding dithiothreitol (DTT) to a final concentration of 1 mM in 10 mM ammonium carbonate, pH 7.5, and incubating at 100 °C for 2 min. At least five spectra were collected for each sample in positive-ion mode, each consisting of signals averaged from 20–100 consecutive laser pulses. Bovine insulin ($[M+H]^+$ = 5734.6 Da) and horse heart cytochrome *c* ($[M+H]^+$ = 12361.1 Da) were used as calibration standards.

Intact MBP-C (approx. 10 pmol/μl in water) was analysed by a similar method except that sinapinic acid (0.7 mg/ml in 70 % acetonitrile containing 0.1 % trifluoroacetic acid) was used as the sample matrix. BSA ($[M+H]^+$ = 66431 Da) was used as the calibration standard. Masses of peptides were calculated from the amino acid sequence of MBP-C using the Protein Abacus software (Lighthouse Data) supplied with the spectrometer.

Subtilisin digestion

Subtilisin digestion of N-terminal tryptic fragments was used to determine the disulphide-bonding pattern of MBP-C. Typically, 0.5 μl of subtilisin (1 pmol/μl) was added to the substrate (approx. 10 pmol/μl) in 10 mM ammonium carbonate, pH 7.5, in a total reaction volume of 5 μl. At various time intervals, 0.5 μl aliquots were removed, applied directly to a MALDI target and mixed with 0.5 μl of α-cyano-4-hydroxycinnamic acid, prepared as described above. Reduced fractions were prepared in the same way except that the reaction was stopped with 1 mM PMSF (added from a 10 mM stock prepared in ethanol), DTT was added (1 mM final concentration) and the sample was heated at 100 °C for 2 min before it was applied to the target.

RESULTS

Overproduction of MBP-C

The quantities of MBP-C that can be purified from rat liver are insufficient for detailed biophysical analysis. The typical yield from a standard preparation is 0.1–0.3 mg from 200 g of rat liver. It is also difficult to eliminate completely contamination from MBP-A [5,25]. These problems have been circumvented by developing an expression system in which MBP-C is overproduced in Chinese-hamster ovary cells [16]. Purification of MBP-C from the culture medium was generally achieved by a single step of affinity chromatography on a mannose–Sepharose column. However, for analytical ultracentrifugation experiments further purification on a MonoS cation-exchange column was required to remove minor contaminants. Using this strategy, 1–2 mg of pure protein was routinely isolated from 400 ml of medium. Amino acid sequencing revealed that the signal sequence is processed correctly.

Figure 1(A) shows samples of MBP-C after each of the purification steps analysed by SDS/PAGE. Under reducing conditions, overproduced MBP-C migrates as a broad band, like native MBP-C, with an apparent molecular mass of 30 kDa. A minor band with an apparent molecular mass of 24 kDa is observed in all preparations (Figure 1A, lane 1). The N-terminal sequence of this contaminant was found to correspond to the sequence of MBP-C beginning at residue Val-54 within the collagenous region. Based on the apparent molecular mass and sequence data, this polypeptide appears to be a proteolytic fragment of MBP-C encompassing the C-terminal half of the collagenous region, the neck and the CRD. This species was largely removed by cation-exchange chromatography (Figure 1A, lane 2).

Comparison of native and overproduced MBP-C

Figure 1(B) shows native and overproduced MBP-C separated by SDS/PAGE under non-reducing conditions. The apparent molecular masses of the three bands (27, 48 and 81 kDa) suggest the presence of monomers and disulphide-linked dimers and

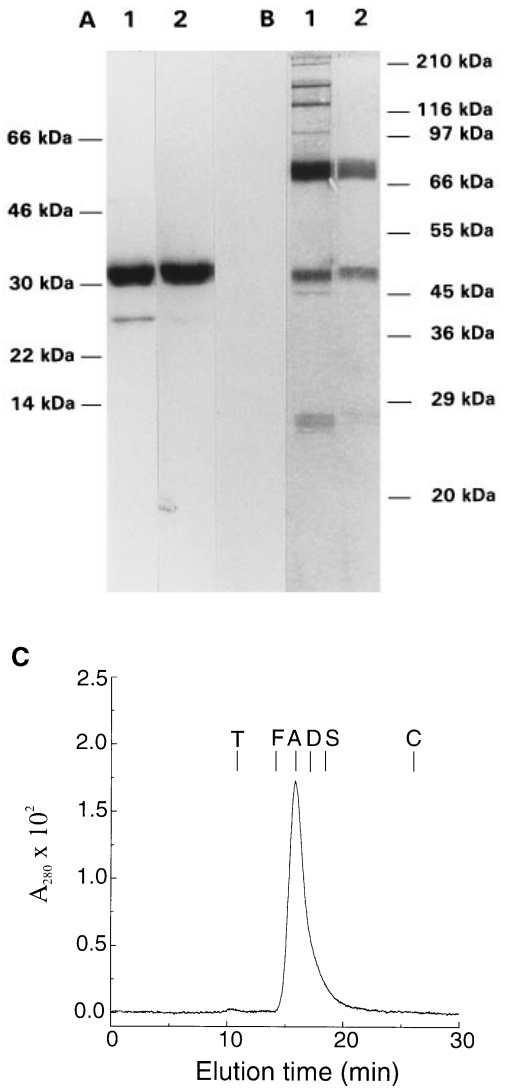

Figure 1 SDS/PAGE and gel-filtration analysis of recombinant MBP-C

(**A**) MBP-C overproduced in Chinese-hamster ovary cells separated on a 17.5% (w/v) gel under reducing conditions. Lane 1, following affinity chromatography on a mannose–Sepharose column; lane 2, following further purification by cation-exchange chromatography. Protein was detected by staining with Coomassie Blue. A proteolytic fragment of MBP-C migrates with an apparent molecular mass of 24 kDa. Molecular-mass markers are shown on the left. (**B**) MBP-C isolated either from rat liver (lane 1) or overproduced in Chinese-hamster ovary cells (lane 2) separated on an SDS/15%-polyacrylamide gel under non-reducing conditions. The overproduced protein in lane 2, repurified by cation-exchange chromatography, corresponds to that shown in (**A**) lane 2. Protein was detected by silver staining. Trimers, dimers and monomers have apparent molecular masses of 81, 48 and 27 kDa respectively. Molecular mass markers are shown on the right. (**C**) Gel-filtration chromatography of recombinant MBP-C (100 μl at 0.5 mg/ml) on a G3000SW gel-filtration column. The elution positions of bovine thyroglobulin (T; 669 kDa), horse spleen apoferritin (F; 443 kDa), sweet-potato β-amylase (A; 200 kDa), yeast alcohol dehydrogenase (D; 150 kDa), BSA (S; 66 kDa) and carbonic anhydrase (C; 29 kDa) are indicated.

trimers. Similar ratios of multimers are observed in the natural and overproduced MBP-C preparations. These ratios are consistent with previous studies [5,6]. The high-molecular-mass bands in the liver sample were identified as MBP-A on immunoblots of parallel gels probed with antibody specific for the CRD of MBP-A [26].

Previous studies have shown that MBP-C elutes from a gel-filtration column with an apparent molecular mass of 194 kDa [6]. Therefore the gel-filtration profiles of recombinant and native preparations were compared to verify that the overproduced protein is assembled correctly. The elution profile of overproduced MBP-C on a TSK G3000SW gel-filtration column is shown in Figure 1(C). The apparent molecular mass was 200 kDa, in agreement with previous data. The Stokes radius was estimated to be 5.0 ± 0.9 nm [19]. SDS/PAGE of fractions across the gel-filtration peak reveal that the multimeric forms (trimer, dimer and monomer) co-migrate under these conditions, suggesting that all are components of higher-order oligomers.

Post-translational modification

Since MBP-C is known to be extensively modified [5,7], it was important to examine the post-translational modifications of the recombinant protein. The broad bands observed on reducing and non-reducing SDS/polyacrylamide gels suggest similar modifications in both native and overproduced preparations. The extent of modification was further examined by MALDI-MS. Under reducing conditions the MALDI-MS profile of overproduced MBP-C is consistent with a species of mass 25295 Da. This value is significantly greater than the calculated mass based on the amino acid sequence (24033.7 Da) [5]. Under non-reducing conditions, peaks consistent with three major species were detected. The observed masses of 25255, 50190 and 75380 Da are compatible with the expected masses of monomer and the masses of disulphide-linked dimer and trimer. Thus the same oligomers are observed as are seen on SDS/polyacrylamide gels.

Post-translational modifications of MBP-C were further analysed by amino acid analysis in conjunction with amino acid sequencing and MALDI-MS of peptide fragments. From amino acid analysis, the composition of 4-hydroxyproline and 5-hydroxylysine was estimated to be 2.3 mol% and 1.6 mol% respectively, reflecting the presence of 5.2 residues of 4-hydroxyproline and 3.6 residues of 5-hydroxylysine per polypeptide chain. Examination of the amino acid sequence of MBP-C indicates that most of the potential sites for hydroxylation must be derivatized, since there are six proline residues and four lysine residues within the putative consensus sequences for hydroxylation: -Xaa-Pro-Gly- and -Xaa-Lys-Gly- [27].

Under both reducing and non-reducing conditions, an additional peak representing a species of mass 19613 Da was detected in preparations of MBP-C before ion-exchange chromatography. This mass is similar to that of the proteolytic fragment identified by SDS/PAGE in these preparations (Figure 1A). The observed m/z ratio by MALDI-MS is in good agreement with the calculated mass of the intact C-terminus of MBP-C (from residues Val-54) assuming complete post-translation modification of the proline and lysine residues (19590.6 Da). Thus the MALDI-MS data are consistent with the N-terminal sequencing data and confirm that the fragment is a C-terminal proteolytic digestion product of MBP-C.

In order to identify individual residues that are modified in MBP-C, the protein was digested with trypsin and the peptide fragments separated by reversed-phase chromatography. In all, 22 major peaks were detected by absorbance at 210 nm, analysed by MALDI-MS and identified based on the amino acid sequence of MBP-C [5]. Identities of fragments derived from the N-terminal cysteine-rich domain and collagenous region were confirmed by amino acid sequencing. Figure 2 summarizes the analysis of the peptides isolated from the 78 N-terminal residues of MBP-C. Peptides Gly^{46}-Lys^{53} and Val^{54}-Arg^{75} elute as single peaks from the reversed-phase column. The MALDI-MS profiles

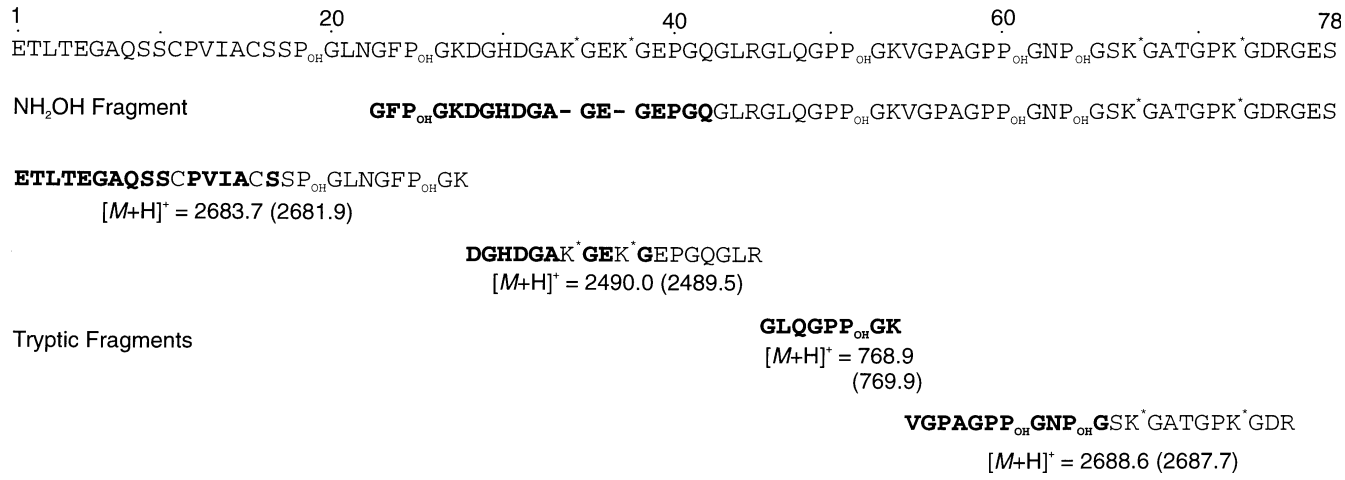

Figure 2 Post-translational modifications of recombinant MBP-C

Peptides were identified by MALDI-MS and the N-terminal sequences were confirmed by amino acid sequencing (bold). The assembled sequence is shown at the top of the Figure. Calculated *m/z* values are shown in parentheses. Removal of the N-terminal 22 amino acids was achieved by cleavage of MBP-C using hydroxylamine. The remaining peptides were generated by trypsin digestion and separated by reversed-phase chromatography. The mass value displayed for Glu[1]-Lys[27] corresponds to the reduced peptide. P_{OH} denotes 4-hydroxyproline and K* denotes glucosylgalactosyl-5-hydroxylysine. Pro-19 and Pro-40 are partially modified. The most abundant peaks, as judged by amino acid sequencing and from the absorbance of peaks resolved by reversed-phase chromatography, are shown.

and sequencing data are consistent with complete hydroxylation of Pro-51, Pro-60 and Pro-63.

Peptides from the N-terminus of MBP-C (Glu[1]-Lys[27]) elute as multiple peaks from the reversed-phase column. The observed masses are consistent with the masses of monomers and disulphide-linked dimers and trimers. However, each oligomer elutes as several partially resolved peaks: four for the trimer; three for the dimer and two for the monomer. Analysis of the two monomer peaks ($[M+H]^+ = 2665.7$ and 2681.6 Da) reveals a mass difference of approx. 16 Da, reflecting incomplete hydroxylation of one or both of the proline residues at positions 19 and 25. Following reduction with DTT, the multiple dimer and trimer peaks also reveal the presence of two forms ($[M+H]^+ =$ 2667.7 and 2683.7 Da) that differ by 16 Da. Under non-reducing conditions, the polypeptides are too large to allow accurate detection of these 16 Da differences, particularly given incomplete resolution of the multiple species by reversed-phase chromatography. Peptide Asp[28]-Arg[45] also elutes as a double peak from the reversed-phase column and the mass difference ($[M+H]^+ =$ 2490.0 and 2505.9 Da) was again 16 Da, suggesting incomplete hydroxylation of Pro-40. Where partial hydroxylation of proline residues was observed, the hydroxylated peptide was found to elute from the reversed-phase column immediately before the unmodified species.

To complement the analysis of tryptic peptides, MBP-C was cleaved with hydroxylamine (between Asn-22 and Gly-23) and the newly exposed N-terminal sequence was determined. Only PTH-4-hydroxyproline was detected at the third cycle, indicating that Pro-25 is fully hydroxylated, and implying that Pro-19 is partially modified. In contrast, only PTH-proline was detected at the 18th cycle, indicating that Pro-40 is largely unmodified. Interestingly, both sites 19 and 40 occur immediately adjacent to breaks in the collagen consensus sequence: Pro-19 is within the first collagen repeat and Pro-40 is immediately before the Gly-Gln-Gly interruption [5]. 3-Hydroxyproline is occasionally found in vertebrate collagens [27]. However, no evidence of 3-hydroxyproline was detected in MBP-C, in spite of the presence

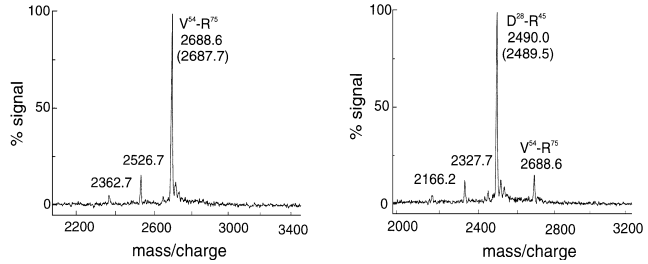

Figure 3 MALDI-MS of glycosylated peptides

In each case the *m/z* is consistent with hydroxylation of lysine and O-linked glycosylation of 5-hydroxylysine residues (Figure 2). Minor peaks correspond to loss of one and two hexose moieties (loss of 162.1 Da). For peptide V[54]-R[75], a single peak was isolated by reversed-phase chromatography. The mass detected by MALDI-MS reflects full hydroxylation of Pro-60 and Pro-63 (Figure 2). Two forms of peptide D[28]-R[45] were resolved by reversed-phase chromatography. MALDI-MS of one form, corresponding to unhydroxylated Pro-40, is shown. Masses of the corresponding peaks for the second form ($[M+H]^+ = 2505.9$, 2343.8 and 2181.8 Da) were 16 Da higher, reflecting hydroxylation of this proline residue.

of two consensus sites for hydroxylation within the sequence -Gly-Pro-4-Hyp- [27].

All four lysine residues within the consensus sequence for hydroxylation were found to be at least partially modified, based on MALDI-MS analysis. In each case, the mass of the peptide is consistent with derivatization of the hydroxylysine residue with galactose or glucosylgalactose, as previously found for MBP isolated from liver [7]. Furthermore, no PTH-amino acid was detected at these positions on sequencing of the peptides. 5-Hydroxylysine glycosides are not usually observed on Edman degradation, whereas the unglycosylated amino acid derivatives can be detected [28]. Figure 3 shows the MALDI-MS profiles of two fractions containing the peptides Asp[28]-Arg[45] and Val[54]-Arg[75]. In each case minor peaks are observed corresponding to the loss of a single hexose moiety (−162.1 Da). It is unclear

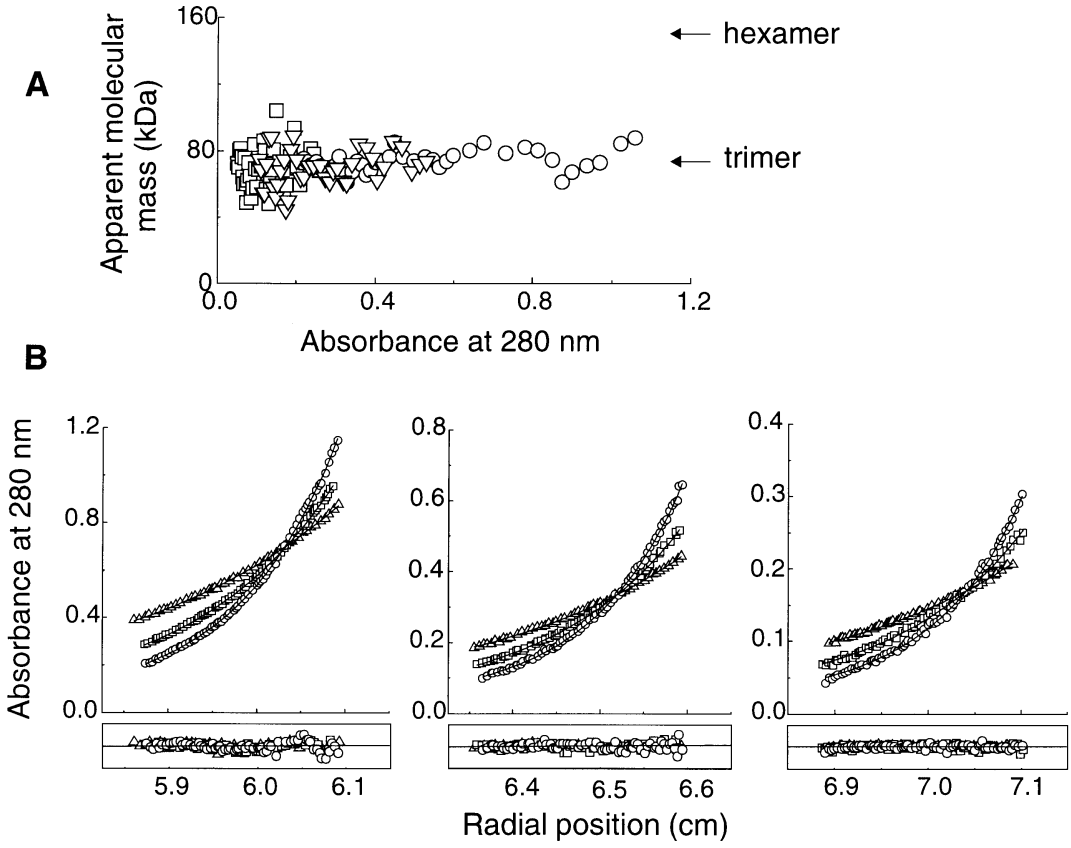

Figure 4 Equilibrium ultracentrifugation of MBP-C

Equilibrium sedimentation of recombinant MBP-C in 50 mM Tris/HCl (pH 7.5)/150 mM NaCl/2.5 mM $CaCl_2$ at 20 °C. (**A**) Apparent weight-averaged molecular mass distribution with concentration for three samples of MBP-C at 12000 rev./min. The loading concentrations were approx. 0.6 mg/ml (○), 0.3 mg/ml (▽) and 0.15 mg/ml (□). The calculated masses of the trimer and hexamer, based on the amino acid sequence of MBP-C, are indicated [5]. (**B**) Equilibrium distributions of three concentrations of MBP-C at three speeds: 8000 (△), 10000 (□) and 12000 (○) rev./min, fitted to a single-species model. The residuals to the fit (± 0.03 A_{280}) are shown at the bottom of the Figure. The loading concentrations were as for (**A**).

whether this result is due to heterogeneity in the peptide samples or is a consequence of the desorption/ionization process.

In summary, analysis of the post-translational modifications of overproduced MBP-C indicates that the collagenous region is extensively modified. The modifications are typical of those found in vertebrate collagens and are consistent with previous studies of the native protein.

Oligomeric state of MBP-C

Based on the apparent molecular mass derived from gel filtration, it has been suggested that MBP-C consists of six polypeptide chains forming two trimeric units [6]. However, MBP-C is likely to be highly asymmetrical because of the presence of an extended triple-helical collagen-like domain. For this reason, analytical ultracentrifugation was used to determine an accurate molecular mass that was independent of the shape of the molecule. Figure 4(A) shows the equilibrium distribution of apparent molecular mass as a function of concentration for three initial loading concentrations of MBP-C. No self-association was observed at protein concentrations up to 12 μM within a range of salt concentrations (0–1.25 M NaCl) and pH (6–7.5), in the presence or absence of Ca^{2+}. Figure 4(B) shows the equilibrium distributions of three loading concentrations of MBP-C at three rotor speeds (8000, 10000 and 12000 rev./min) fitted to a single-species model. The weight-averaged molecular mass was estimated to be 77000 ± 2000 Da from three independent experiments, which corresponds closely to the mass of a trimer of the modified polypeptide (approx. 75900 Da from MALDI-MS).

Sedimentation of MBP-C was measured in 50 mM Tris, pH 7.5, containing 150 mM NaCl and 2.5 mM $CaCl_2$. The sedimentation coefficient ($s^0_{20,w}$) was 3.28 ± 0.16 S from three independent experiments. Using this value and the weight-averaged molecular mass from the equilibrium experiment, the frictional ratio and Stokes radius of MBP-C were calculated to be 2.07 ± 0.15 and 5.8 ± 0.5 nm respectively. The Stokes radius is in good agreement with the value estimated by gel filtration. The molecular mass of MBP-C was also calculated from the diffusion coefficient ($D = 3.88 \times 10^{-7}$ cm$^2 \cdot$s^{-1}), which was determined by light scattering, and the sedimentation coefficient. The estimated value of 74000 ± 4000 Da is very similar to that determined by equilibrium ultracentrifugation and supports the contention that MBP-C is trimeric.

The biophysical data for MBP-C are consistent with a highly asymmetrical molecule, as would be expected from its domain organization. The collagenous region contains 19 Gly-Xaa-Yaa repeats. Based on the crystal structure of a collagen-like peptide, the average unit height for a single repeat is 8.4 Å [29]. Combining this value with the measurements of the trimeric head structure [12], the overall dimensions would be 8 nm maximum diameter for the globular head groups and an overall length of approx. 25 nm.

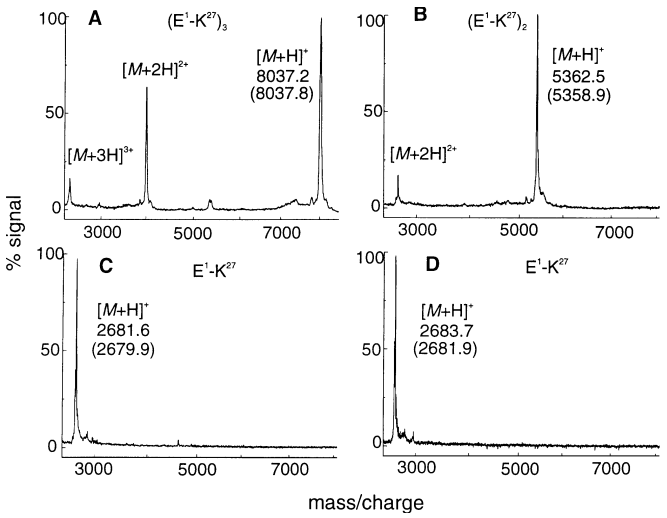

Figure 5 Disulphide-linked peptides from the N-terminus of MBP-C, examined by MALDI-MS

(**A**) Trimer; (**B**) dimer; (**C**) monomer. In each case the *m/z* observed corresponds closely to the value calculated from the amino acid sequence (in parentheses). The calculated value indicated assumes full hydroxylation of Pro-19 and Pro-25, although partial modification of Pro-19 identified by a difference of 16 Da was observed in some fractions. (**D**) Reduced trimer. Both the trimer (shown) and the dimer could be reduced to monomers with 1 mM DTT.

Figure 6 Two possible isomers of the trimeric N-terminal peptide of MBP-C

The interchain disulphide bonds can theoretically be arranged in either an asymmetrical (isomer 1) or a symmetrical pattern (isomer 2).

Interchain disulphide bonding

MBP-C contains six cysteine residues, four of which are highly conserved within the CRD and form two intrachain disulphide bonds: Cys^{133}-Cys^{222} and Cys^{200}-Cys^{214} [30]. The two remaining cysteine residues are located at the N-terminus of the protein at positions 11 and 16 and must be responsible for interchain disulphide bonds between the polypeptides. Analysis of MBP-C isolated from both liver and Chinese hamster ovary cells by SDS/PAGE and MALDI-MS indicates that significant quantities of dimer and monomer are present in the samples. Similar ratios were observed in all preparations. However, all of the MBP-C behaves as trimers by sedimentation ultracentrifugation and elutes as a single peak on gel filtration. Therefore at least two forms must exist: a fully interchain disulphide-linked trimer, which appears to be the most abundant species, and a trimer comprising disulphide-linked dimer plus an additional non-covalently associated subunit. The presence of a possible third form, consisting of three non-covalently associated monomers, cannot be ruled out entirely. However, the ratio of dimers to monomers seen on gels is close to 2:1 (Figure 1B), so the abundance of such a species, if present, must be very low.

In order to establish which cysteine residues are involved in disulphide bonds, alkylation of free thiol groups within MBP-C with iodo[2-^{3}H]acetic acid in the presence and absence of the denaturant guanidine hydrochloride was attempted [31]. No labelling of protein was detected following SDS/PAGE. It was concluded that there are no free thiol groups within the oligomeric forms of MBP-C and both cysteine residues in each polypeptide chain must form disulphide bonds. MALDI-MS analysis of tryptic fragments corresponding to the disulphide-linked trimer and dimer and the monomer yielded masses that correspond closely to the calculated masses from the amino acid sequence (Figures 5A–5C). Thus the N-terminal cysteines are not modified by linkage to free cysteine or glutathione, as has been observed in certain secreted proteins [32]. Both the trimeric and dimeric peptides could be reduced by the addition of DTT, confirming that the polypeptide chains are linked by disulphide bonds (Figure 5D).

Two alternative isomers can be considered for the covalent trimer, each with three interchain disulphide bonds (Figure 6). The bonding pattern is either asymmetrical (isomer 1) or symmetrical (isomer 2). Two conformations are also possible for the non-covalent trimer where one polypeptide chain contains an intrachain bond and the other two chains are linked by Cys^{11}-Cys^{16} and Cys^{11}-Cys^{16} or Cys^{11}-Cys^{11} and Cys^{16}-Cys^{16} disulphide bonds. N-terminal amino acid sequencing could be used to analyse the bonding pattern in the tryptic peptides, although it was not possible to quantify cystine at each cycle, since di(PTH) cystine is relatively unstable. Edman degradation of the trimer revealed di(PTH)cystine in cycles 11 and 16, reflecting the presence of isomer 1, since its release would be expected only at cycle 16 if isomer 2 were the exclusive form. Di(PTH)cystine was also detected in the 11th cycle during Edman degradation of the dimer indicating the presence of the Cys^{11}-Cys^{11} and Cys^{16}-Cys^{16} conformation.

The suggested interchain disulphide bonding patterns in MBP-C were confirmed by cleavage of the N-terminal tryptic fragments between Ala-15 and Cys-16 with subtilisin. Analysis of the cleavage products of the trimeric peptide by MALDI-MS clearly establishes the presence of isomer 1 in Figure 6. Peaks were observed corresponding to all three dimeric peptides: (Glu^1-Ala^{15})/(Cys^{16}-Lys^{27}), (Glu^1-Ala^{15})$_2$ and (Cys^{16}-Lys^{27})$_2$. Peaks were also detected corresponding to the loss of dimers from the trimeric substrate: (Glu^1-Lys^{27})/(Glu^1-Ala^{15})$_2$ and (Glu^1-Lys^{27})/(Glu^1-Ala^{15})/(Cys^{16}-Lys^{27}). Disulphide bond cleavage can occur at low frequency in MALDI-MS [33], and peaks corresponding to monomers were observed. However, no peaks were detected that were consistent with the loss of single chains (Glu^1-Ala^{15} or Cys^{16}-Lys^{27}) from the substrate. Thus the digestion patterns are consistent with the closed, disulphide-bonded isomer 1 in Figure 6. The identities of peptides were confirmed by reducing the disulphide bonds in the peptide mixture with DTT (Figure 7B).

Subtilisin digestion was also used to confirm the arrangement of disulphides in the dimeric peptide. Although peaks corresponding to homodimers (Glu^1-Ala^{15})$_2$ and (Cys^{16}-Lys^{27})$_2$ were

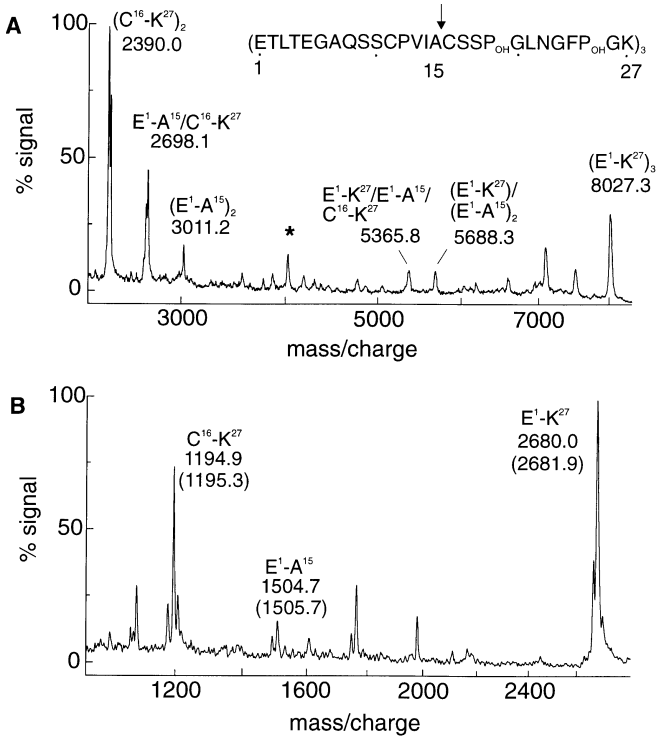

Figure 7 Arrangement of interchain disulphide bonds in the trimeric peptide determined by MALDI-MS

(**A**) Subtilisin digest of the trimeric N-terminal tryptic peptide. The sequence of the substrate and the cleavage position between the cysteine residues are shown in the inset. The *m/z* values of the peptides correspond closely to the values calculated from the sequence: 2387.6, 2698.0, 3008.3, 5376.9, 5687.2 and 8037.8 Da. While the smaller fragments show clear evidence of partial hydroxylation of the proline residues (double peaks separated by 16 Da), the individual species of the larger peptides could not be resolved. The calculated masses indicated in the Figure assume full modification. Only the masses of those species that distinguish the disulphide-bonding pattern are shown. The peak marked with an asterisk (*) corresponds to doubly-charged substrate. Other commonly observed peaks are consistent with the loss of Gly^{23}-Lys^{27} (−502.6 Da), Glu^{1}-Ala^{7} (−701.7 Da), Glu^{1}-Ser^{9} (−916.9 Da) and Ser^{17}-Lys^{27} (−1074.2 Da) from one or more polypeptide chains. (**B**) Digest from a similar time point as in (**A**) following reduction with 1 mM DTT. The *m/z* values of digestion products are shown together with the value calculated from the amino acid sequence (in parentheses).

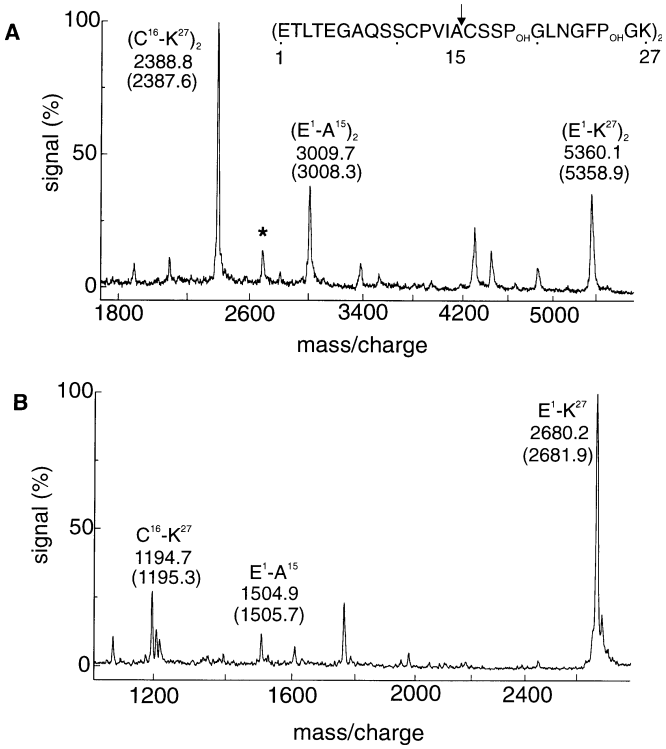

Figure 8 Subtilisin digestion of the N-terminal dimeric peptide analysed by MALDI-MS

Digests were carried out as described in Figure 7. The sequence of the substrate and the cleavage position between the cysteine residues are shown in the inset. (**A**) Profile under non-reducing conditions. Although peaks are detected for the homodimers at *m/z* 2388.8 and 3009.7, consistent with the trimer digest, no fragment was detected corresponding to the heterodimer Glu^{1}-Ala^{15}/Cys^{16}-Lys^{27} (calculated mass = 2698.0 Da), indicating that the disulphide-bonding pattern is Cys^{11}-Cys^{11} and Cys^{16}-Cys^{16}. The peak marked with an asterisk (*) corresponds to doubly-charged substrate. Peaks not labelled are consistent with the loss of Gly^{23}-Lys^{27} (−502.6 Da), Glu^{1}-Ala^{7} (−701.7 Da), Glu^{1}-Ser^{9} (−917.0 Da) and Ser^{17}-Lys^{27} (−1074.2 Da) from one or both polypeptide chains. The calculated masses of peptides are shown in parentheses. (**B**) Digest from a similar time point as (**A**) following reduction with 1 mM DTT.

observed (Figure 8A), there was no evidence of the heterodimer (Glu^{1}-Ala^{15})/(Cys^{16}-Lys^{27}), which could be detected in digests of the trimeric peptide (Figure 7A). Substrate was chosen in which the proline residues were fully hydroxylated (identified by MALDI-MS) to ensure that the heterodimer (Glu^{1}-Ala^{15})/(Cys^{16}-Lys^{27}) (calculated $m/z = 2698.0$ Da) could be distinguished from doubly-charged substrate ($m/z = 2679.9$). The identities of the peptides were again confirmed by reduction with DTT (Figure 8B). All of the digestion products observed for the dimer (as for the trimer) are consistent with fully disulphide-linked peptides. Attempts to confirm the presence of an interchain disulphide bond in the monomer by proteolysis were not successful as none of the proteases tested cleaves between the cysteines.

The observation that MBP-C polypeptides can be cross-linked into two different configurations may seem surprising. 'Artefactual' formation of the non-covalent trimer during purification can never be completely ruled out. However, this possibility seems unlikely since formation of a disulphide-linked dimer and monomer from the covalent trimer would require the breakage and formation of at least two disulphide bonds. No evidence of disulphide shuffling was observed in MBP-C or any of the proteolytic peptides on storage. In addition, analysis of the dimeric and monomeric peptides indicates that they adopt specific conformations (dimer Cys^{11}-Cys^{11} and Cys^{16}-Cys^{16}; monomer Cys^{11}-Cys^{16}) that are not random, as would be expected following reorganization of the disulphide bonds.

DISCUSSION

Oligomeric structure of MBP-C and other collectins

The biophysical analysis reported here indicates that MBP-C adopts a highly extended conformation in solution. This finding is consistent with images of other collectins observed by rotary-shadowing electron microscopy [8,11]. Images of human MBP, for example, reveal globular heads connected by central stalks of length 13.2 ± 2.0 nm [8]. The electron micrographs are compatible with the crystal structure of a trimeric fragment of MBP-A, composed of the CRD and neck region, assuming that most of the length of each stalk consists of a collagen-like structure and the globular heads correspond to clusters of three CRDs [12].

Many collectins form large oligomeric structures based on a trimeric unit. For example, serum MBPs, including human MBP and rat MBP-A, form multimers containing up to 18 polypeptide

chains [8]. However, the biophysical data reported here indicate that MBP-C is a single trimer despite sharing substantial sequence similarity with MBP-A [5]. Collectin-43, another member of the collectin family, has also been reported to be a simple trimer, based on electron microscopic and biophysical studies [11]. Whereas MBP-C shares sequence identity with collectins that adopt the bouquet-like conformation, collectin-43 has a longer collagenous region, containing 38 Gly-Xaa-Yaa repeats, and has a sequence similar to SP-D and conglutinin, both of which form cruciform structures. Since MBP-C and collectin-43 are only distantly related [34], it can be concluded that the similar oligomeric structures that they adopt in solution arose through distinct evolutionary routes.

The arrangement of two cysteine residues separated by four amino acid residues at the N-terminal junction of the collagenous region is found in many of the collectins, including SP-D, conglutinin and collectin-43 [35]. A similar pattern is found in rat MBP-A and human MBP, which each contain one extra N-terminal cysteine residue. It seems likely that the asymmetrical interchain disulphide-bonding arrangement found in MBP-C is conserved in other members of the family. Additional disulphide bonds formed through the extra cysteine residue in the serum MBPs probably link polypeptides in separate trimers to form the larger covalent oligomers observed on non-reducing gels [8]. Thus the analysis of MBP-C reported here forms a useful starting point for visualizing the organization of the more complex serum homologues.

Asymmetry at the N-terminus of MBP-C

The results reported here indicate that the polypeptide chains in MBP-C must adopt an asymmetrical conformation at each end of the collagenous domain. This asymmetry arises because the constituent chains of a collagen triple helix are staggered [29]. As a consequence of this stagger, residues that connect the neck to the C-terminal end of the collagenous domain must be arranged differently in the three protomers that make up the trimer. As illustrated in Figure 9, the trimeric coiled-coil of α-helices in the neck of MBP-C has three-fold rotational symmetry, while the collagen triple helix has a three-fold screw axis, so that the overall molecule cannot have simple rotational symmetry. Similarly, from the chemical evidence presented here, residues at the N-terminal end of the collagenous domain must also be arranged differently in the three copies. The alignment of the polypeptides introduced by the collagen triple helix must orientate the cysteine residues at the N-terminus of the protein to allow formation of the asymmetrical disulphide bonding pattern observed.

In the case of fibrillar collagens, it is believed that the polypeptides are modified as they enter the endoplasmic reticulum and that they subsequently assemble in a C-terminal-to-N-terminal direction. Initiation of assembly is thought to result from folding of C-terminal propeptides [27]. A fragment comprising the neck region and the CRD of MBP-C can associate to form a trimer, in the absence of the collagenous domain [37]. Thus the CRD and neck regions in collectins may perform a similar function to the C-terminal propeptides in fibrillar collagens by acting as the nucleation site for assembly of the collagen triple helix. Formation of an asymmetrical disulphide bonding pattern at the N-terminus would require previous formation of the collagen triple helix. Therefore it could be envisaged that assembly of MBP-C occurs in a C-terminal-to-N-terminal direction, the neck region initiating formation of the collagen triple helix, which in turn determines the structure of the N-terminal cysteine-rich domain.

A further consequence of the stagger in the polypeptide chains of the collagen triple helix is that there are multiple possible conformations for both the covalent and the non-covalent trimeric forms of MBP-C. The different conformers arise because the three chains are not spatially equivalent at the junction of the N-terminal cysteine-rich domain and collagenous region. Since the disulphide bonds of the covalent trimer are arranged in an asymmetrical pattern, six conformers are possible. For example, the disulphide bond between Cys^{11}-Cys^{11} might occur between chains 1 and 2, 1 and 3 or 2 and 3. There are also three possible conformers for the non-covalent trimer.

The arrangement of cysteine residues in the N-terminal domain of the collectins is similar to that found in certain non-fibrillar collagens, the fibril-associated collagens with interrupted triple helix, which consist of collagenous regions separated by non-helical (NC) domains [38]. The C-terminal end of each of these molecules contains a collagenous region (COL1) followed by an NC1 domain. The order of the two domains is reversed compared with that found in the collectins, and the COL1–NC1 junction contains two cysteine residues separated by four amino acid residues. NMR studies reveal that, following air oxidation, synthetic peptides corresponding to the COL1–NC1 junction of type XIV collagen form a trimer linked by three interchain bonds, in which the pattern of the disulphide bonds is asymmetrical [39]. The NMR data are consistent with two conformers in which one of the chains folds into a turn-like arrangement linked by two disulphide bonds to the other chains. By analogy it seems likely that MBP-C will adopt only certain specific conformations from the nine possible arrangements arising as a result of the stagger in the collagen triple helix.

Possible flexible regions within collectins

The asymmetry observed in the present studies suggests that there may be multiple regions of flexibility in MBP-C. It has previously been noted that MBP-A and -C each contain a single interruption in the Gly-Xaa-Yaa repeat pattern within the collagenous region. Calculations from electron micrographs of human MBP are consistent with this region forming a kink where the trimeric stems angle away from the core [8]. Although MBP-C is a single trimer, this region may still be flexible and introduce a bend into the protein. The junction between the collagenous region and the neck represents a second potentially flexible region (Figure 9). Human class A macrophage scavenger receptors also contain α-helical coiled-coil domains adjacent to collagenous regions (although in the opposite orientation). In electron micrographs of these proteins the angle between the two domains ranges from 0 to 180° indicating that the junction is highly flexible [40]. Since the MBP-C polypeptides are out of alignment, due to the stagger introduced by the collagen triple helix, the α-helix–collagen junction may also be flexible in MBP-C and other members of the collectin family.

Role of MBP-C

The clustered binding sites in MBP-A are well suited for binding very-high-density mannose- or *N*-acetylglucosamine-containing structures such as those found on fungal and bacterial surfaces, although the sugar-binding sites are too far apart for high-affinity multivalent interactions with a typical mammalian high-mannose oligosaccharide [12]. The binding specificity of MBP-C for monosaccharides is similar to MBP-A [30], which suggests that its role may be associated with host defence. However, significant differences are observed in binding to more complex ligands, implying that MBP-C may interact with the tri-mannose core of N-linked oligosaccharides found on glycoproteins, whereas

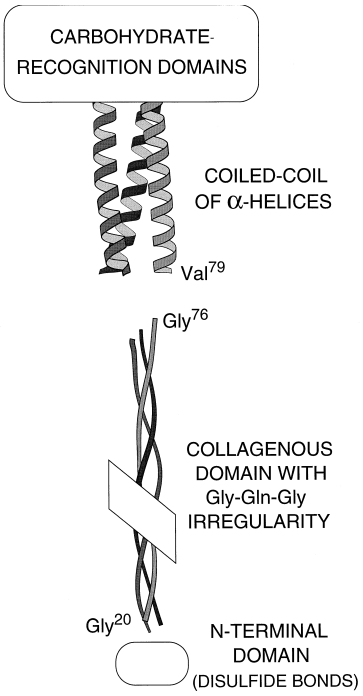

Figure 9 Modelled structure of the N-terminal portion of MBP-C

The coiled-coil of the neck region has been modelled using the co-ordinates for MBP-A (Brookhaven Protein Database entry 1rtm), and the collagenous domain is based on a collagen-like synthetic peptide (Brookhaven Protein Database entry 1cag). Only the beginning and the end of the collagenous segment is shown. The Figure was prepared using MOLSCRIPT [36].

MBP-A does not [37]. This difference has been attributed to the presence of a secondary sugar-binding site in the CRD of MBP-C that is not detected in MBP-A [41]. The ability to interact with endogenous glycoproteins, combined with the fact that MBP-C is not highly oligomeric, suggests that its retention in the liver may result from interactions with specific high-affinity ligands. Thus despite their similarities at the sequence level and in domain organization, MBP-C and MBP-A could have distinct biological functions.

Roles for MBP-C in cell–cell interactions or glycoprotein trafficking in the liver have been suggested [25]. The possibility that MBP-C might cross-link cells to each other or to the matrix would be consistent with the earlier proposal that it forms dimers of trimers, since the carbohydrate-binding sites in such a dimer could be disposed at opposite ends of the molecule. However, the current demonstration that MBP-C consists of a single trimer is not consistent with this idea of homotypic binding together of glycosylated surfaces. If MBP-C does, in fact, act as a bridge, it would probably be between carbohydrates bound by the C-terminal CRDs and some other type of structure that interacts with the N-terminal cysteine-rich or collagenous domains. The binding of lipoproteins to the collagenous domain of the macrophage scavenger receptor [42] and the binding of other collectins to cell-surface collectin/C1q receptors [43] suggest possible modes of interaction.

Received 20 January 1997/28 February 1997; accepted 5 March 1997

Funding for this work was provided by the Wellcome Trust. We thank David Harvey and Bernard Küster for their advice concerning the MALDI-MS experiments.

REFERENCES

1 Ikeda, K., Sannoh, T., Kawasaki, N., Kawasaki, T. and Yamashina, I. (1987) J. Biol. Chem. **262**, 7451–7454
2 Kuhlman, M., Joiner, K. and Ezekowitz, R. A. B. (1989) J. Exp. Med. **169**, 1733–1745
3 Super, M., Levinsky, R. J. and Turner, M. W. (1990) Clin. Exp. Immunol. **79**, 144–150
4 Drickamer, K. and Taylor, M. E. (1993) Annu. Rev. Cell Biol. **9**, 237–264
5 Drickamer, K., Dordal, M. S. and Reynolds, L. (1986) J. Biol. Chem. **261**, 1034–1046
6 Mizuno, Y., Kozutsumi, Y., Kawasaki, T. and Yamashina, I. (1980) J. Biol. Chem. **256**, 4247–4252
7 Colley, K. J. and Baenziger, J. U. (1987) J. Biol. Chem. **262**, 10290–10295
8 Lu, J., Thiel, S., Wiedemann, H., Timpl, R. and Reid, K. B. M. (1990) J. Immunol. **144**, 2287–2294
9 Persson, A., Chang, D., Rust, K., Moxley, M., Longmore, W. and Crouch, E. (1989) Biochemistry **28**, 6361–6367
10 Yokota, Y., Arai, T. and Kawasaki, T. (1995) J. Biochem. (Tokyo) **117**, 414–419
11 Holmskov, U., Laursen, S. B., Malhotra, R., Wiedemann, H., Timpl, R., Stuart, G. R., Tornoe, I., Madsen, P. S., Reid, K. B. M. and Jensenius, J. C. (1995) Biochem. J. **305**, 889–896
12 Weis, W. I. and Drickamer, K. (1994) Structure **2**, 1227–1240
13 Heinrikson, R. L. and Meredith, S. C. (1984) Anal. Biochem. **136**, 65–74
14 Laemmli, U. K. (1970) Nature (London) **227**, 680–685
15 Sambrook, J., Fritsch, E. F. and Manatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16 Kaufman, R. J., Davies, M. V., Wasley, L. C. and Michnick, D. (1991) Nucleic Acids Res. **19**, 4485–4490
17 Kaufman, R. J. (1991) Methods Enzymol. **185**, 537–566
18 Graham, F. L. and van der Eb, A. J. (1973) Virology **52**, 456–467
19 Laurent, T. C. and Killander, J. (1964) J. Chromatogr. **14**, 317–330
20 Sober, H. A. (1970) Handbook of Biochemistry: Selected Data For Molecular Biology, The Chemical Rubber Co, Cleveland, OH
21 Johnson, M. L., Correia, J. A., Yphantis, D. A. and Halvorson, H. R. (1981) Biophys. J. **36**, 575–588
22 Cohn, E. J. and Edsall, J. T. (1943) Proteins, Amino Acids And Peptides as Ions and Dipolar Ions, pp. 370–381, Reinhold, New York
23 Bornstein, P. and Balian, G. (1977) Methods Enzymol. **47**, 132–145
24 Matsudaira, P. (1987) J. Biol. Chem. **262**, 10035–10038
25 Mori, K., Kawasaki, T. and Yamashina, I. (1988) Arch. Biochem. Biophys. **264**, 647–656
26 Quesenberry, M. S. and Drickamer, K. (1992) J. Biol. Chem. **267**, 10831–10841
27 Kadler, K. E. (1994) Protein Profile **1**, 525–534
28 Butler, W. T. (1982) Methods Enzymol. **82**, 339–346
29 Bella, J., Easton, M., Brodsky, B. and Berman, H. M. (1994) Science **266**, 75–81
30 Ng, K. K.-S., Drickamer, K. and Weis, W. I. (1996) J. Biol. Chem. **271**, 663–674
31 Liu, T.-Y. (1977) in The Proteins, (Neurath, H. and Hill, R. L., eds.), pp. 240–402, Academic Press, New York
32 Reddy, P., Sparvoli, S., Fagioli, C., Fassina, G. and Sitia, R. (1996) EMBO J. **15**, 2077–2085
33 Crimmins, D. L., Saylor, M., Rush, J. and Thoma, R. S. (1995) Anal. Biochem. **226**, 255–361
34 Drickamer, K. (1993) Curr. Opin. Struct. Biol. **3**, 393–400
35 Hoppe, H.-J. and Reid, K. B. M. (1994) Protein Sci. **3**, 1143–1158
36 Kraulis, P. J. (1991) J. Appl. Crystallogr. **24**, 946–950
37 Childs, R. A., Feizi, T., Yuen, C. T., Drickamer, K. and Quesenberry, M. S. (1990) J. Biol. Chem. **265**, 20770–20777
38 Mayne, R. and Brewton, R. G. (1993) Curr. Opin. Cell Biol. **5**, 883–890
39 Lesage, A., Penin, F., Geourjon, C., Marion, D. and van der Rest, M. (1996) Biochemistry **35**, 9647–9660
40 Resnick, D., Chatterton, J. E., Schwartz, K., Slayter, H. and Krieger, M. (1996) J. Biol. Chem. **271**, 26924–26930
41 Quesenberry, M. S., Lee, R. T. and Lee, Y. C. (1997) Biochemistry **36**, 2724–2732
42 Rohrer, L., Freeman, M., Kodama, T., Penman, M. and Krieger, M. (1990) Nature (London) **343**, 570–572
43 Peerschke, E. I. B., Malhotra, R., Ghebrehiwet, B., Reid, K. B. M. and Sim, R. B. (1993) J. Leukocyte Biol. **53**, 179–184