Physical Characterization and Crystallization of the Carbohydrate-recognition Domain of a Mannose-binding Protein from Rat

William I. Weiss‡§, Gregg V. Cricchio, H. M. Krishna Murthy‡¶, Wayne A. Hendrickson‡, and Kurt Drickamer†

From the Department of Biochemistry and Molecular Biophysics and the Howard Hughes Medical Institute, Columbia University, New York, New York 10032

A portion of rat mannose-binding protein A (MBP-A), a Ca\(^{2+}\)-dependent animal lectin, has been overproduced in a bacterial expression system, biochemically characterized, and crystallized. A fragment corresponding to the COOH-terminal 116 residues of native MBP-A, produced by subtilisin digestion of the bacterially expressed protein, contains the carbohydrate-recognition domain (CRD). Gel filtration, chemical cross-linking, and crystallographic self-rotation function analyses indicate that the subtilisin fragment is a dimer, although the complete bacterially expressed fragment, containing the neck and CRD of MBP-A, is a trimer. Crystals of the minimal CRD, obtained only as a complex with a Man\(_3\)GlcNAc\(_2\)Asn glycopeptide, diffract to Bragg spacings of at least 1.7 A. Several trivalent lanthanide ions (Ln\(^{3+}\)) can substitute for Ca\(^{2+}\), as assessed by their ability to support carbohydrate binding and to protect the CRD from proteolysis in a manner similar to that observed for Ca\(^{2+}\). These assays indicate that Ln\(^{3+}\) binds about 30 times more tightly than Ca\(^{2+}\) to the CRD, and that two Ca\(^{2+}\) or Ln\(^{3+}\) bind to each monomer, a result confirmed by determination of the Ho\(^{3+}\) positions in a Ho\(^{3+}\)-containing crystal of the CRD. Crystals grown in the presence of Ln\(^{3+}\) belong to different space groups from those obtained with Ca\(^{2+}\) and are therefore not useable for traditional crystallographic phase determination methods, but are well-suited for high resolution structure determination by multiwavelength anomalous dispersion phasing.

Mammalian mannose-binding proteins (MBPs) found in serum and liver recognize a variety of saccharide structures, and mediate defensive reactions (1–7). It is believed that the MBP-mediated recognition of carbohydrate structures, such as those found on various pathogenic bacteria and yeasts, may serve to target defense reactions to these organisms. This hypothesis is supported by the intrinsic properties of the MBPs in vitro, since they can act as opsonins (8) and mediate fixation of complement (9). Increased levels of MBP mRNA observed during the acute phase response suggest that levels of the human serum MBP may rise when exposure to such pathogens occurs (10). The correlation of reduced levels of serum MBP with a propensity for repeated, severe bacterial infections in certain individuals is also consistent with the suggestion that the MBPs mediate important protective functions (11).

The primary structures of the two rat MBPs (MBP-A and MBP-C) (12) and one human homolog (13) indicate that they are composed of several discrete domains. Following a short, NH\(_2\)-terminal, cysteine-rich segment, each MBP consists of a collagenous domain containing 18–20 Gly-X-Y repeats, a neck region, and a COOH-terminal globular domain (Fig. 1). The presence of collagen-like domains both in complement protein C1q and in MBP-A suggests that their shared ability to fix complement may be mediated by this domain. Digestion of MBPs with collagenase produces a fragment consisting of the COOH-terminal 154 amino acids, which retains carbohydrate-binding activity (12). This portion constitutes a carbohydrate-recognition domain (CRD) (14).

The COOH-terminal carbohydrate-recognition domains of MBPs show sequence homology with the CRDs found in a number of other animal lectins that require Ca\(^{2+}\) and neutral pH for carbohydrate binding. These proteins have been termed C-type lectins (14). Other members of this family include the mammalian and avian asialoglycoprotein receptors (also known as hepatic lectins), LEC-CAMS such as lymphocyte homing receptors, and pulmonary surfactant apoproteins. These proteins contain C-type CRDs attached to effector domains that vary with the function of the molecule. C-type CRDs are characterized by a sequence motif of 32 conserved amino acids at fixed intervals, including two disulfide bonds, distributed over approximately 120 amino acids. The conserved residues appear to form a general Ca\(^{2+}\)-dependent carbohydrate-binding framework; specificity for particular carbohydrates is presumably encoded in some of the nonconserved residues, as different C-type lectins show a wide range of distinct carbohydrate specificities.

In an attempt to understand the molecular basis of Ca\(^{2+}\)-dependent carbohydrate binding and specificity of C-type lectins, as well as to define the three-dimensional basis for the sequence motif shared among these proteins, we have initiated physical studies of several CRDs ultimately aimed at obtaining their atomic level structures, including complexes with various ligands. In this report, we describe the definition of the minimal carbohydrate-binding domain for rat MBP-A using bacterial expression and proteolytic methods. We have...
prepared crystals of a complex between this domain and a Man\(\beta\)GlcNAc2Asn glycopeptide ligand in the presence of Ca\(^{2+}\), and we have also prepared crystals of the domain in the presence of various lanthanide ions for use in crystallographic phase determination. Using biochemical and crystallographic techniques, we have investigated the oligomeric state of various fragments of MBP-A, and the binding of cations to the CRD.

**EXPERIMENTAL PROCEDURES**

**Materials**—Lanthanide trichloride hexahydrates were purchased from Aldrich Chemical Co. and used without further purification. Lanthanide stock solutions of 0.5-1 m were made to the nominal concentration; no correction was made for excess water, despite the extremely hygroscopic character of these compounds. Man\(\beta\)GlcNAc2Asn isolated from ovalbumin, with the following structure, was purchased from Biocarb Chemicals, Lund, Sweden.

Man\(\alpha\)1--6Man\(\alpha\)1--3Man\(\beta\)1--4GlcNAc2--5GalNAcβ1--3GlcNAcβ1--2Man\(\beta\)1--4GlcNAcβ1--3Man\(\alpha\)1--2Man

**Guaniidine HCl and subtilisin** were obtained from Boehriinger-Mannheim Biochemicals. Bis(sulfosuccinimidyl)suberate was purchased from Pierce Chemical Co. Polyethylene glycol and proteins for gel filtration chromatography standards were from Sigma.

**Protein Production**—Expression was performed in *E. coli* strain JA221 bearing the pNIIImopA2-derived expression plasmid (15). Medium LB (16) (5 liters) was inoculated with 100 ml of a stationary culture, and grown for 2.5 h. Following addition of isopropyl-1-thio-\(\beta\)-D-galactopyranoside to a final concentration of 2 mM, cells were grown for an additional 2 h, harvested, and washed with 10 mM Tris chloride, pH 7.8. Lysis was accomplished by sonication (6 bursts of 30 s duration) in 100 ml of the wash buffer, and an insoluble pellet was recovered by centrifugation at 10,000 \(\times\) g for 15 min. This pellet was dissolved in 100 ml of 6 M guanidine hydrochloride containing 100 mM Tris chloride, pH 7.0, and 0.01% 2-mercaptoethanol by brief sonication and gentle mixing for 30 min. Insoluble material was removed by centrifugation at 100,000 \(\times\) g for 30 min. The supernatant solution was diluted with 4 volumes of loading buffer (1.25 M NaCl, 25 mM Tris chloride, pH 7.8, 25 mM CaCl\(_2\)) and dialyzed against 20 volumes of loading buffer, with five changes over 2 days. Insoluble material was again removed by centrifugation at 100,000 \(\times\) g for 60 min, and the supernatant solution was applied to a 10-ml column of mannose-Sepharose (17) equilibrated in loading buffer. After application of sample, the column was washed with 100 ml of loading buffer and eluted with elution buffer (1.25 M NaCl, 25 mM Tris chloride, pH 7.8, 2.5 mM EDTA). Fractions containing the expressed CRD were identified by SDS-polyacrylamide gel electrophoresis of aliquots following the method of Laemmli (18). Yields of 20-50 mg of protein were obtained from 5 liters of starting culture.

For further purification, fractions containing MBP-A-F1' were dialyzed against 10 mM NaCl, 1 mM Tris chloride, pH 7.8, lyophilized, and dissolved in 0.01-0.1 \(\times\) their original volume of H\(_2\)O. Aliquots (100 \(\mu\)l) of this material were applied to a 10 \(\times\) 30-cm C3 high performance liquid chromatography column (Beckman Ultrapore) equilibrated with 10% acetonitrile, 0.1% trifluoroacetic acid. The column was eluted with a linear gradient in which the acetonitrile concentration was increased at a rate of 1% min. Absorbance was monitored at 280 nm. Under these conditions, the expressed fragment elutes at approximately 35% acetonitrile. Fractions containing pure MBP-A-F1' were lyophilized and then dissolved in water at a concentration of 25-50 mg/ml (assayed by the method of Bradford (19), using ovalbumin as standard). This material was diluted into appropriate buffers for various experiments.

**Proteolytic Digestion**—For proteolysis experiments, MBP-A-F1' eluted from the mannosese-Sepharose column was adjusted to 25 mM CaCl\(_2\). Tests were performed by digestion with various concentrations of subtilisin, followed by gel electrophoresis. Conditions which gave maximal yield of fragment MBP-A-F1' were: 25 \(\mu\)g subtilisin final concentration, for 1 h at 37°C were then used to digest 30-50 \(\mu\)l of pooled affinity column eluate containing 100-200 mg of MBP-A-F1'. After digestion, the reaction mixture was applied directly to a 5-ml column of mannose-Sepharose at 4°C. The column was rinsed with 50 ml of loading buffer and eluted as above. Fractions containing MBP-A-F2, identified by gel electrophoresis, were pooled, dialyzed, and purified by reverse phase chromatography as described above.

For analytical digestions, lyophilized MBP-A-F2 was dissolved at a concentration of 0.5 mg/ml in 50 mM Tris chloride, pH 7.8. Various amounts of CaCl\(_2\) or HoCl\(_2\) were added, followed by subtilisin at a concentration of 3 \(\mu\)g/ml. Digestion (15 min at 37°C) was terminated by addition of double-strength gel sample buffer containing 20 mM EDTA and 1% 2-mercaptoethanol followed by immediate heating to 100°C for 5 min. Following gel electrophoresis and staining with Coomassie Blue, protease-resistant material was quantified by scanning on a Bio-Rad 620 video densitometer. Data were analyzed using a nonlinear least squares fitting program (SigmaPlot, Jandel Scientific, Corte Madera, CA).

**Chemical Cross-linking**—MBP-A fragments were dissolved at a concentration of 0.1 mg/ml in 0.1 M sodium HEPES, pH 7.5, containing 10 mM CaCl\(_2\). Aliquots (25 \(\mu\)l) were treated with bis(sulfosuccinimidyl)suberate for 1 h at 22°C. Samples were diluted 1:1 with double-strength gel sample buffer containing 2% 2-mercaptoethanol and heated to 100°C for 5 min.

**Gel Filtration Chromatography**—Molecular weights were estimated by gel filtration chromatography on a G3000SW column (0.8 \(\times\) 30-cm GlassFak, LKB Instruments). Elution was performed at a flow rate of 0.2 ml/min with a buffer containing 250 mM Tris acetate, pH 7.8, 25 mM NaCl, and 25 mM CaCl\(_2\). Five sets of 50 ml of 1 \(\mu\)g of standards (Table I) and 8 \(\mu\)g of MBP-A preparations were injected. Standards for calibration were yeast alcohol dehydrogenase (150 kDa), ovalbumin (43 kDa), bovine erythrocyte carbonic anhydrase (29 kDa), and myoglobin (17 kDa).

**Ligand Binding Assay**—Iodination of the CRD of MBP-A and binding to invertase-coated nitrocellulose were performed as previously described (20).

**Amino Acid and Carbohydrate Content Analysis of Crystals**—For each analysis, a single crystal was harvested from a 1-ml hanger drop to which approximately 10 \(\mu\)l of reservoir solution had been added as a 100 \(\mu\)l aliquot. 20-50 \(\mu\)l of reservoir solution with a 0.5-mm diameter glass capillary, taking care to minimize the transfer of excess liquid with the crystal. The crystal was gently moved through this solution for about 3 min to rinse any free protein and carbohydrate present in the original hanging drop solution from its surface. After transfer of the crystal to another 200 \(\mu\)l of reservoir solution, this rinsing procedure was repeated. Next, the crystal was transferred to a 25% (w/v) solution of polyethylene glycol 3350, adjusted to approximately pH 8 with NaOH, and the rinsing procedure repeated to remove residual Tris buffer from the crystal. After repeating this step, the crystal was transferred to 75% PEG 3350 in reservoir solution and allowed to equilibrate for 1 h at 104°C in 5.7 N HCl. In test experiments, these conditions were found to give nearly complete hydrolysis of the peptide, with minimal destruction of amino sugar (21). Amino acid and glucosamine analysis was performed on an Applied Biosystems 420A amino acid analyzer.

**X-ray Diffraction Measurements and Calculations**—X-ray diffraction data from a single monoclinic crystal were collected on a multiwire area detector (San Diego Multiwire Systems) with a Rigaku AFC6 gonostat. The x-ray source was graphite-monochromated CuK\(_\alpha\) radiation produced by a Rigaku RU-200 rotating anode generator operating at 50 kW \(\times\) 100 mA with a 0.5-mm fine-focus cathode. A total of 400° of data were collected in 0.15° frames, and integrated on line using the San Diego software. Redundant Lorentz polarization corrected intensity measurements were merged in 5° batches to a unique set using the ROTA VATA and AGROVATA programs in the CCP4 package (22). Data were collected at 20°C. Data collection statistics are given in Table I.

Diffraction data from a single orthorhombic HoK\(_\beta\)O\(_\gamma\) derivative crystal were collected in 0.10° frames on the spherical drift chamber detector on the D23 beam line at the LURE synchrotron, Orsay, France (23). Three data sets were collected at nominal wavelengths of 1.5356, 1.5363, and 1.5485 \(\AA\) as part of a multiwavelength anomalous dispersion (MAD) phasing experiment. In order to measure Bijvoet-related reflections simultaneously, 106° of data were collected by rotating about the crystallographic 6 axis, and 45° about the a axis. Data were collected at 4°C. Lorentz polarization corrected

---

integrated intensities were obtained with the MADNES package (24), followed by the profile fitting algorithm of Kabsch (25). These data were binned into 2" batches and reduced to a unique data set with ROTAVATA and a modified version of AGROVATA which rejects outlying measurements only with respect to redundant copies of the same Friedel mate I(+h) or I(-h), rather than with respect to the mean of all measurements of I(±h). Statistics for the 1.4400 A data are shown in Table I.

Self-rotation function calculations were performed using the fast rotation function method (26) as implemented in the MERLOT package (27), and the Patterson search method (28) implemented in the X-PLOR program (29).

RESULTS

Biochemical Characterization of MBP-A Fragments

Definition of the Carbohydrate-binding Domain within MBP-A—The quantities of MBPs that can be isolated from serum or liver are limited, and resolution of the two types of MBP is necessary in order to study homogeneous proteins. To circumvent these problems, a bacterial expression system was developed to produce in large quantities, the protein accumulates in in-

expression vector remain at the NH, terminus of the fragment. When this portion of the molecule retains the ability to bind mannose-containing ligands in a Ca\(^{2+}\)-dependent manner, it was used as the starting point for these studies.

Limited digestion of MBP-A-F1' with clostripain in the presence of Ca\(^{2+}\) resulted in removal of the extra amino acids derived from the expression vector. The predominant fragment produced consisted of residues 73 through 221 of MBP-A (MBP-A-F1; see Fig. 1). However, it was contaminated by a slightly longer fragment starting at residue 69, and it proved difficult to obtain a homogeneous preparation of fragment MBP-A-F1. In contrast, digestion of MBP-A-F1' with subtilisin in the presence of Ca\(^{2+}\) produces a homogeneous protease-resistant core (Fig. 2). Similar results have been obtained using trypsin (data not shown). The protease-resistant fragment, designated MBP-A-F2, can be repurified by affinity chromatography on mannose-Sepharose, demonstrating that it contains the CRD. The affinity-isolated subfragment was further purified by reverse phase chromatography and subjected to automated Edman degradation for 20 cycles. The results revealed an NH2-terminal sequence corresponding to the sequence of intact MBP-A starting at residue 107. This indicates that the carbohydrate-binding activity resides in the COOH-terminal 115 amino acids of MBP-A that comprise the fragment MBP-A-F2 (Fig. 1).

Oligomeric Structure of MBP-A Fragments—Natural MBP-A is an oligomer of the constituent polypeptide, with a molecular weight of approximately 650,000. The higher oligomers appear to be formed by association of trimeric building blocks which are believed to be stabilized, at least in part, by the formation of collagen-like triple helices in the NH2-terminal portions of the constituent polypeptide (12, 31, 32). Since formation of the larger native complexes may involve interactions between several different portions of the polypeptide, it was of interest to determine which interactions are preserved in the fragments that have been produced in the bacterial system. This question was investigated by gel filtration chromatography and chemical cross-linking methods. As shown in Fig. 3A, fragment MBP-A-F1' can be cross-linked to form dimers and trimers. This fragment elutes with an apparent molecular weight of 65,000 on a gel filtration column (Fig. 4A), consistent with the size expected of a trimer. These results suggest that the COOH-terminal portion of MBP-A retains the ability to form trimers, even in the absence of the collagen-like domain.

The behavior of the fragment MBP-A-F2 contrasts sharply with that of MBP-A-F1'. In a cross-linking experiment (Fig. 3B), dimers are the major oligomers observed. A limited amount of trimer formation is seen, but this band appears at the position expected for a trimer of MBP-A-F1'; faint extra...
bands corresponding to MBP-A-F1' monomers and dimers are also observed. The trimer band thus appears to be due to a small amount of undigested MBP-A-F1' that is not separated from MBP-A-F2 by the high performance liquid chromatography purification. MBP-A-F2 elutes with an apparent molecular weight of 24,000, approximately the expected size of a dimer. The self-rotation function analysis of two different crystal forms presented below also supports the conclusion that MBP-A-F2 is a dimer.

Crystallization of Ca\(^{2+}\)-containing MBP-A Fragments

**MBP-A-F1'**—Attempts to crystallize MBP-A-F1' produced several crystal forms, all of which grew only in the presence of Ca\(^{2+}\) and Man\(_6\)GlcNAc\(_2\)Asn. This oligosaccharide has been shown to bind to preparations of MBP-A-F1' with an estimated dissociation constant of 1.3 mM (33). Growth of crystals was critically dependent on the ligand concentration. One form was produced by the hanging drop vapor diffusion method in which a reservoir of 15% polyethylene glycol (PEG) 3350 or 8000, 20 mM CaCl\(_2\), 100 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.02% NaN\(_3\) with drops formed by mixing 0.75–1.5 µl of 20 mg/ml of MBP-A-F1', 10 mM CaCl\(_2\), 5 mM NaCl, 1.1 mM Man\(_6\)GlcNAc\(_2\)Asn, with an equal volume of reservoir (i.e. starting concentration of 15 mM CaCl\(_2\) in the drop). Parallel experiments in which the CaCl\(_2\) or carbohydrate ligand were replaced with water produced no crystals. Crystals appear as early as 18 h after setting up the drop, and grow to as large as 0.4 × 0.3 × 0.2 mm\(^3\) in 3–7 days. The crystals always grow from drops containing a heavy precipitate. Fig. 5 shows a typical crystal. The space group is monoclinic, P2\(_1\), (Table I). The crystals often start as small roughly rectangular prisms, with two opposing edges cut off to produce an irregular hexagonal cross-section; full size crystals display variable morphologies, but generally have the hexagonal cross-section. The crystallographic b axis appears to be parallel to one of the shorter edges of the hexagon, such that it emerges from the large face of the crystal. Diffraction corresponding to Bragg spacings of at least 1.7 Å has been observed on synchrotron sources.

Amino acid and glucosamine analysis performed on a washed crystal revealed the presence of ~1 carbohydrate

Crystallization of a Carbohydrate Recognition Domain

**Fig. 3. Chemical cross-linking of MBP-A-F1' and MBP-A-F2.** Aliquots (25 µl) of binding domains (A, MBP-A-F1'; B, MBP-A-F2) were treated with 1, 0 mM; 2, 0.16 mM; 3, 0.3 mM; 4, 0.6 mM; 5, 1.2 mM; 6, 2.5 mM; 7, 5 mM; 8, 10 mM bis(sulfosuccinimidyl)suberate for 1 h at 22 °C. Following electrophoresis on a SDS-polyacrylamide gel (15%), cross-linked complexes were detected by staining with Coomassie Blue.

**Fig. 4. Gel filtration of MBP-A fragments.** Samples were purified by reverse phase chromatography and analyzed on a G3000SW column. A, 1 µg of MBP-A-F1'; B, 1 µg of MBP-A-F2. Positions of standards are indicated at top: AlcDH, alcohol dehydrogenase; BSA, bovine serum albumin; OVAL, ovalbumin; CA, carbonic anhydrase; Mb, myoglobin; APROT, aprotinin.

These crystals grow as plates thickened at one end to give a wedge-shaped morphology. They diffract weakly (typical Bragg spacing limit of about 4 Å), and display high mosaic spread when the face of the crystal is perpendicular to the x-ray beam, and split spots when the face is oriented parallel to the beam. Precession photographs of any zone show no symmetry other than Friedel symmetry. The triclinic P1 unit cell can be described by parameters a = 65.9 Å, b = 68.1 Å, c = 138.1 Å, α = 96.0°, β = 103.8°, γ = 90.1°. The c axis lies roughly perpendicular to the face of the crystal. Assuming partial specific volumes of 0.73 cm\(^3\)/g for protein and 0.65 cm\(^3\)/g for carbohydrate, such a cell would contain 2–5 trimers for a reasonable solvent content (i.e. 40–76% solvent). Two other crystal forms were obtained using PEG as the precipitant, but at lower pH and different salt mixtures. They display weak diffraction and high mosaic spread similar to the triclinic crystals, and they likewise have multiple copies of the trimer in their asymmetric units, based on calculations similar to those described above.

**MBP-A-F2**—The proteolytic susceptibility of the NH\(_2\)-terminal region of MBP-A-F1' suggested that this region is rather flexible and thus may interfere with the formation of well-ordered crystals, and that MBP-A-F2 would be a more favorable candidate for crystallization. Crystals of MBP-A-F2 were grown by equilibrating a reservoir of 15–20% PEG 3350 or 8000, 20 mM CaCl\(_2\), 100 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.02% NaN\(_3\) with drops formed by mixing 0.75–1.5 µl of 20 mg/ml of MBP-A-F1', 10 mM CaCl\(_2\), 5 mM NaCl, 1.1 mM Man\(_6\)GlcNAc\(_2\)Asn, with an equal volume of reservoir (i.e. starting concentration of 15 mM CaCl\(_2\) in the drop). Parallel experiments in which the CaCl\(_2\) or carbohydrate ligand were replaced with water produced no crystals. Crystals appear as early as 18 h after setting up the drop, and grow to as large as 0.4 × 0.3 × 0.2 mm\(^3\) in 3–7 days. The crystals always grow from drops containing a heavy precipitate. Fig. 5 shows a typical crystal. The space group is monoclinic, P2\(_1\), (Table I). The crystals often start as small roughly rectangular prisms, with two opposing edges cut off to produce an irregular hexagonal cross-section; full size crystals display variable morphologies, but generally have the hexagonal cross-section. The crystallographic b axis appears to be parallel to one of the shorter edges of the hexagon, such that it emerges from the largest face of the crystal. Diffraction corresponding to Bragg spacings of at least 1.7 Å has been observed on synchrotron sources.

Amino acid and glucosamine analysis performed on a washed crystal revealed the presence of ~1 carbohydrate
Crystallization of a Carbohydrate Recognition Domain

15 mM Ca²⁺  15 mM Ho³⁺  1 mM Ho³⁺

FIG. 5. Crystals of MBP-A-F2. Scale bars represent 0.1 mm. All crystals shown were grown in the presence of Man₆GlcNAc₂Asn. Left, monoclinic crystal grown in the presence of Ca²⁺, using a starting concentration of 15 mM CaCl₂ in the drop. Center, tetragonal crystal grown in the presence of Ho³⁺, using a starting concentration of 15 mM HoCl₂ in the drop. Right, orthorhombic crystal grown in the presence of Ho³⁺, using a starting concentration of 15 mM HoCl₂ in the drop.

TABLE I

<table>
<thead>
<tr>
<th>MBP-A-F2 crystallographic data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting cation concentration in drop</td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>15 mM Ca²⁺</td>
</tr>
<tr>
<td>15 mM Ln³⁺</td>
</tr>
<tr>
<td>1 mM Ln³⁺</td>
</tr>
</tbody>
</table>

Data collection statistics

| Crystal                      | Bragg spacing limit Rmax No. unique reflections Rarch Centric Acentric |
|------------------------------|-------------------------------------------------|-----------------|---------------------|
| Monoclinic                   | 2.8 0.062 6276 (93)                              | 0.018 0.092     |
| Orthorhombic                 | 2.8 0.030 5621 (97)                              | 0.018 0.092     |

* R = \( \sum |I(h)| - |\langle I(h)\rangle|\) / | | \langle I(h)\rangle| |, where I(h) is the ith measurement of I(h) and \( \langle I(h)\rangle\) is the weighted mean of all measurements of I(h).

Characterization of Tetragonal Ln³⁺-containing Crystals—Ln³⁺-substituted crystals were obtained by equilibrating reservoirs of 16–22% PEG 3350, 30 mM NiCl₂, 100 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.02 mM NaN₃ with a hanging drop made by adding 0.75–1.5 μl of 20 mg/ml MBP-A-F2, 5 mM NiCl₂, 1.1 mM Man,GlcNAc₂Asn to an equal volume of reservoir (i.e. starting concentration of 15 mM NiCl₂ in the drop). Long rod-shaped crystals with a square cross-section were obtained in Gd²⁺, Tb³⁺, Dy³⁺, Ho²⁺, and Tm³⁺, with typical sizes 0.4 × 0.05–0.08 × 0.05–0.08 mm² (Fig. 5). Subsequent experiments showed that these crystals will grow in the absence of ligand. In addition to their different morphologies, these crystals show much weaker birefringence than the native crystals, and diffract at most to 2.7 Å Bragg spacings on still photographs. These crystals are in the tetragonal space group P₄₋₂,₂,₂ (Table I), as determined from screened precession photographs of crystals grown in the presence of TbCl₃ and ligand. The crystallographic c axis coincides with the long direction of the crystal, and the rods are bounded by [110] and (001) faces. If the partial specific volume of the protein is 0.73 cm³/g, the most likely solvent content is 59%, with 4 monomers, or 2 dimers, per asymmetric unit, assuming multiples of dimers.

Characterization of Orthorhombic Ln³⁺-containing Crystals—The change in space group accompanying the substitution of Ca²⁺ with Ln³⁺ suggested that in addition to substituting for the Ca²⁺, the Ln³⁺ was somehow changing the packing of the molecule by, for example, forming bridges between molecules in the crystal lattice, thereby altering packing...
Crystallization of a Carbohydrate Recognition Domain

Monoclinic
$\chi = 180^\circ$

Orthorhombic
$\chi = 180^\circ$

**FIG. 6. Self-rotation function maps.** Results of self-rotation function calculations are given in the spherical polar angle convention of Rossmann and Blow (41), where $\psi$ and $\phi$ describe the orientation of the rotation axis with respect to the crystallographic unit cell, and $\chi$ is the rotation about this axis. The $\chi = 180^\circ$ section is plotted as an equiareal polar projection. $\psi$ is displayed radially from 0 to 90°, with circles of constant $\psi$ shown at 30, 60, and 90°; $\phi$ runs counterclockwise from the crystallographic $a$ axis, with lines of constant $\phi$ shown every 45°. Left, monoclinic crystal, contoured in intervals of 0.5 standard deviations ($\sigma$), starting at 2.5 $\sigma$. The peak at $\psi = 77^\circ$, $\phi = 42^\circ$ is 51% of the origin. All data observed between 8.0 and 2.8 Å were used in the calculation. Right, orthorhombic crystal, contoured in intervals of 0.25 $\sigma$, starting at 1.25 $\sigma$. The peak at $\psi = 81^\circ$, $\phi = 39^\circ$ is 24% of the origin. Finer sampling of this rotation function reveals the maximum to be at $\psi = 80.5^\circ$, $\phi = 39.6^\circ$, $\chi = 178.5^\circ$. All centric and all acentric Friedel-paired data (see text) observed between 8.0 and 2.8 Å were used in the calculation. Both plots were made with a maximum Patterson integration radius of 20 Å. In both cases, the peaks shown here are seen over a wide range of maximum Patterson integration radii (12-25 Å), provided that the high resolution data cutoff is at least 4 Å.

**FIG. 7. Cation dependence of ligand binding and protease resistance of MBP-A-F2.** Binding to invertase-coated filters (upper panels) and protease resistance (lower panels) as a function of Ca$^{2+}$ (A, B) or Ho$^{3+}$ (C, D) concentration were measured as described under "Experimental Procedures." The data have been fit with an equation that assumes second order dependence on cation concentration (solid curves), which in each case provided the best fit to the observed values. For example, fitting the data for Ca$^{2+}$ dependence of ligand binding to first, second, and third order equations resulted in summed squared residuals of 0.103, 0.018, and 0.092, respectively.
and Yb\(^{3+}\), with Ho\(^{3+}\) giving the largest crystals most reliably.

A somewhat lower polyethylene glycol concentration of 11-15% appears to be optimal for the growth of this crystal form. Subsequent experiments showed that these crystals will grow in the absence of ligand, and no carbohydrate was detected in an amino acid and glucosamine composition analysis of a washed crystal (Table II).

The 1 mM Ln\(^{3+}\) crystals have a morphology distinct from the other two: they grow as elongated rods, sometimes with bevelled edges, and their ends are often invaginated (Fig. 5). Single crystals typically have sizes of about 0.25-0.40 × 0.05-0.10 × 0.05-0.10 mm\(^3\); larger crystals are almost always mechanically twinned, with two or more crystals grown into one another. As in the monoclinic case, these crystals grow in drops containing a heavy precipitate. Diffraction to 1.5-Å Bragg spacings has been observed on a synchrotron source. Precession photographs of crystals grown from starting concentrations of 1 mM TbCl\(_3\), HoCl\(_3\), or YbCl\(_3\) show that the space group is orthorhombic, P2\(_1\)2\(_1\)2\(_1\) (Table I), with the b axis lying in the long direction of the crystals, and the a and c axes perpendicular to the broad faces of the rod. Calculation of the solvent content for these crystals gives 70% solvent by volume for 1 MBP-A-F2 monomer in the asymmetric unit, 41% for 2, and 11% for 3. By the same arguments given for the monoclinic crystals, the asymmetric unit most likely contains 2 monomers. To confirm this, a self-rotation function was run on the 1.4400 Å data set (see "Experimental Procedures") in which the structure factor amplitudes of acentric reflections were set equal to the average of the two Bijvoet mates to minimize the effect of the anomalous dispersion signal (unpaired acentric measurements were discarded in this procedure). The self-rotation map shows an unambiguous peak at \(\chi = 180^\circ\) (Fig. 6), confirming the presence of two copies in the asymmetric unit.

To determine the stoichiometry of Ho\(^{3+}\) binding, an anomalous difference Patterson map (36) was computed from the orthorhombic 1.4400 Å diffraction data (Fig. 8). Solution of the Patterson revealed four Ho\(^{3+}\) sites per asymmetric unit. After refinement of these sites against the observed Bijvoet differences (\(R = 0.318\) for \(F(h) > 4\sigma_F\) and \(|\Delta F| > 2\sigma_F\), where \(\Delta F = |F(+h)| - |F(-h)|\)), difference Fourier maps were inspected for the presence of any other Ho atoms; none were found. Moreover, a Patterson map calculated from the refined Ho model reproduced all significant peaks in the Bijvoet difference Patterson map. The four sites are composed of two pairs of Ho atoms, with the two atoms of a pair separated by 8.5 Å. A self-rotation function computed from structure factors calculated from the refined Ho\(^{3+}\) positions revealed the same peak as that found in the rotation function made from the observed protein structure factors, indicating that the orthorhombic crystals, although devoid of ligand, should contain functional MBP-A-F2. The presence of the monoclinic and orthorhombic crystal forms contain a dimer in the asymmetric unit, consistent with the solution evidence that this fragment is a dimer. Furthermore, the monoclinic form appears to contain a specific complex between MBP-A-F2 and a high-mannose oligosaccharide, although the formal possibility remains that the presence of stoichiometric amounts of this oligosaccharide is not due to specific binding, but is merely a fortuitous requirement for packing of the crystal lattice. Finally, the solution data showing functional substitution of Ca\(^{2+}\) by Ln\(^{3+}\) demonstrate that the orthorhombic crystals, although devoid of ligand, should contain functional MBP-A-F2. The presence of the carbohydrate ligand in the Ca\(^{2+}\)-containing ( monoclinic) crystals, and its absence in the Ln\(^{3+}\)-containing ( orthorhombic) crystals does not have an obvious explanation, since Ln\(^{3+}\) supports ligand binding by MBP-A (Fig. 7). A definitive explanation of this difference awaits determination of the structures of both crystal forms.

1 M. S. Quesenberry and K. Drickamer, manuscript in preparation.
Although the orthorhombic crystal form of MBP-A-F2 is not useful as an isomorphous derivative for traditional phase determination methods such as multiple isomorphous replacement, it is ideally suited for analysis from MAD measurements.

The MAD phasing method does not require isomorphous crystals since phase information is obtained by observing the variation of anomalous dispersion from a single crystalline specimen as a function of x-ray wavelength. The method does require the presence of a sufficiently strong anomalous scattering signal to permit accurate measurement of such variation. The lanthanides display extremely large anomalous dispersion effects at their L\textsubscript{III} edges (39). The substantial anomalous signal (Table I) and the facile solution of the anomalous difference Patterson at a wavelength remote from the Ho absorption edge (Fig. 8), as well as the low R-value for the refinement of the Ho\textsuperscript{++} positions, indicates that a large anomalous dispersion signal is in fact present in the diffraction data from these crystals. The successful determination of the orthorhombic crystal structure by MAD phasing will be reported elsewhere.\textsuperscript{2}

Previous studies have correlated the Ca\textsuperscript{2+} dependence of protease sensitivity of C-type CRDs with their ability to bind carbohydrate ligands. In the case of the subtilisin fragment of the chicken receptor that mediates endocytosis of glycoproteins containing terminal N-acetylgalactosamine, the Ca\textsuperscript{2+} dependence of ligand binding and protease sensitivity are virtually indistinguishable at neutral pH, with an apparent Ca\textsuperscript{2+} dissociation constant of about 2 mM (35). The Ca\textsuperscript{2+} dependence of the MBP-A CRD, as assessed by ligand binding and proteolytic susceptibility, closely resembles that of the chicken receptor. In both cases, the concentration dependence is best fit by assuming 2 Ca\textsuperscript{2+}/monomer. Results of similar studies in the rat asialoglycoprotein receptor have also been interpreted as indicating that there are two Ca\textsuperscript{2+}-binding sites per CRD (40). This suggests that these CRDs ligate divalent cations in a similar manner, and it is noteworthy that a number of the conserved amino acids in the C-type CRD motif are potential Ca\textsuperscript{2+} ligands (i.e. Asp, Glu, Asn, Gln) (14). These data, along with the crystallographically determined stoichiometry presented here, make it seem likely that the presence of two Ca\textsuperscript{2+} binding sites is a common characteristic of C-type CRDs.

An apparent Ca\textsuperscript{2+} dissociation constant of 98 \textmu M has been determined for the transition from nonbinding to saccharide-binding states of the unresolved rat MBPs (5), and similar values have been obtained for two separated human MBPs (7). The higher dissociation constant obtained here for MBP-A-F2 may indicate that regions of the MBP polypeptide outside of the CRD have an effect on ligand binding or stability of the CRD, either by interactions of the CRD with other portions of the same polypeptide, or with other polypeptides in the intact oligomer.

The oligomeric structure of the two fragments MBP-A-F1 and MBP-A-F2 presents an intriguing problem, as dimers and trimers must contain different intermonomer interactions. The susceptibility of intact MBPs to mammalian collagenase, which digests triple-stranded collagenous regions specifically, indicates that a trimer stabilized by a collagen-like coiled-coil is the building block of the higher order oligomers observed in gel filtration measurements (37). The results presented here suggest that portions of the neck that connects the CRD to the collagenous domain, which is present in the larger fragment but absent from the smaller, contributes to stabilization of the trimer in the intact molecule. The 7 extra NH\textsubscript{2}-terminal amino acids in this fragment introduced in the expression construct do not appear to be responsible for trimer formation, as digestion of this fragment with clostripain removes these residues but leaves a trimer.\textsuperscript{4} The dimmerization of MBP-A-F2 is more difficult to rationalize. The formation of dimers may reflect interactions between portions of the CRD that contribute to the higher order oligomerization of trimers. Alternatively, dimerization may be due to interactions not present in the natural molecule, but which become accessible when the rest of the polypeptide is removed. In the only other CRD for which the oligomeric structure has been investigated, that of the chicken hepatic lectin, the equivalent fragment is a monomer (35). Thus, the formation of dimers observed in MBP-A-F2 does not appear to be a general feature of C-type CRDs.

Acknowledgments—We thank Nydia Ruiz for assistance in protein purification, Maureen Taylor for advice on cross-linking experiments, and Richard Kahn and Roger Fourme for participation in data collection at the LURE synchrotron, and we appreciate support from LURE.

REFERENCES

22. CCP4 (1979) The SERC (UK) Collaborative Computing Project No. 4, Daresbury Laboratory, Warrington WA4 4AD, United Kingdom
20686 Crystallization of a Carbohydrate Recognition Domain