

A HUMAN MANNOSE-BINDING PROTEIN IS AN ACUTE-PHASE REACTANT THAT SHARES SEQUENCE HOMOMOLOGY WITH OTHER VERTEBRATE LECTINS

BY R. ALAN B. EZEKOWITZ, LINDA E. DAY, AND GARY A. HERMAN

From the Division of Hematology/Oncology, The Children's Hospital and Dana-Farber Cancer Institute, Harvard Department of Pediatrics, Boston, Massachusetts 02115

Vertebrates possess proteins capable of mannose oligosaccharide recognition (1). An integral membrane protein involved in clearance of mannose terminal glycoproteins was first described on Kupffer cells in rat liver (2). This mannose receptor is a 175-kD integral membrane glycoprotein (3) that is expressed on the surface of mature macrophages in man, rabbits, guinea pigs, and rodents (4). The expression of the receptor on macrophages mediates clearance of mannose terminal glycoproteins and, more importantly, binding and uptake of pathogens with mannose-rich cell walls (5, 6). Another group of mannose lectin-like proteins has been isolated from the livers of rats (7–10) and man (11, 12) and subsequently found in the serum of rats (7, 8), rabbits (13), and man (11, 14, 15). These proteins form part of a family of lectin-like proteins that include five membrane-bound hepatic receptors (reviewed in reference 16). The amino acid homology has been extended (17–20) and also includes two nonvertebrate proteins: a humoral lectin isolated from the fresh fruit fly, *Sarcophaga perigina* (21), and a lectin isolated from the coelomic fluid of the sea urchin, *Anthocidaris crassispina* (22). Other proteins capable of binding specific sugars like the pentraxins (23) are not related to this larger family of lectin-like proteins.

Much more is known about the structure of mannose-binding proteins than is known about its function. Two homologous rat mannose-binding proteins (MBPs),¹ rat MBP-A and rat MBP-C, are the best characterized (24). They exist as at least trimers of a 32-kD monomer that can be divided into three domains: an NH₂-terminal domain rich in cysteine that mediates interchain disulfide bonds, a collagen domain, and a COOH-terminal domain that is the putative carbohydrate-binding region (24). The mannose-binding proteins isolated from various mammalian sera have a molecular mass of ~350 kD consisting of multimers (15, 25) of the above described monomer. In rat sera, the MBP-A is the predominant form and is able upon binding to mannan-coated erythrocytes to activate complement via the classical pathway (25). The rat MBP-C, a 200-kD multimer (24, 25), is only found at low levels in rat serum despite the presence of a signal peptide (24). In addition, the rat MBP-C is not able to activate complement (25).

While the rat seems to possess distinct liver and serum MBPs, the presence of

This work was supported by National Institutes of Health grant ROI AI-23786-OI (to R. A. B. Ezekowitz). Address correspondence to R. A. B. Ezekowitz.

¹ Abbreviation used in this paper: MBP, mannose-binding protein.

similar distinct forms in the human has not yet been clearly documented (11, 12, 14, 15). Mannose-binding proteins isolated from human liver and serum have identical apparent molecular weights, binding characteristics, and share immunological cross-reactivity (11, 12, 14, 15). As structural heterogeneity of homologous rat MBPs appears to mediate functional differences, we wished to extend our knowledge of the structure of human MBPs in order to elucidate its function.

In this study we report the isolation of cDNA clones that encode for a human mannose-binding protein of 30 kD. The encoded protein is most homologous to the rat MBP-C, but it also bears close homology to the rat MBP-A and other vertebrate lectins. In addition, this protein's synthesis in the liver is induced as part of the acute-phase response and, unlike the rat MBP-C, is the predominant serum form in humans.

Materials and Methods

Isolation of cDNA and Genomic Clones. 2,000–3,000 colonies per nitrocellulose filter of a human liver cDNA library constructed in a plasmid pKT218 were plated as described (26). Duplicate filters were prehybridized in a solution containing 0.75 M NaCl, 50 mM sodium phosphate, pH 7.4, 5 mM EDTA; 5× Denhardt's solution, and 0.1% SDS for 1 h at 45°C. A rat MBP-C cDNA was digested with Xho I and Eco RI (24), gel purified on a 1% low melt agarose gel, radiolabeled by random priming, and used as a probe at 10⁶ cpm/ml. This fragment contains the sequence encoding the COOH-terminal carbohydrate-binding domain. Filters were hybridized overnight at 42°C and washed at 45°C in 2× SSC for 30 min and then in 1× SSC for a further 30 min. Positive clones were plaque purified. In addition, a HEPG2 λ gt10 cDNA library, a gift from Dr. D. Kwiatkowski (Massachusetts General Hospital, Boston, MA) was plated in an *Escherichia coli* host C600 as described (27) and screened with a radiolabeled human mannose-binding protein cDNA.

A human genomic library, prepared by partial digestion of genomic DNA with Mbo I and ligation into the Bam HI site of EMBL 3A, was plated in *E. coli* strain LE 392 and screened at high stringency with radiolabeled human MBP cDNA. Filters were subjected to stringent wash in 0.1× SSC at 68°C. Relevant portions of the genomic clone of the human MBP were identified by Southern blot analysis of restriction enzyme digests of phage DNA. An Eco RI fragment that hybridized to the 5' end of the cDNA clone was subcloned into the Eco RI site of plasmid pUC13. An Eco RI/Bam HI digestion of the subcloned fragment was resolved on a 1% agarose gel, yielding two fragments of 600 and 400 bp. These fragments were vectorally cloned into Bam HI/Eco RI-digested M13 mp18 and mp19, and the nucleotide sequence of both strands was determined. The smaller fragment sequence overlapped with the available 5' cDNA sequence and extended further in the 5' direction.

DNA Sequence Analysis. Restriction fragments of the cDNA and genomic clones were subcloned into M13 mp18 or M13 mp19 and sequenced by the dideoxynucleotide chain termination method (28) using ³⁵S-labeled dCTP (29). Each region was sequenced on both strands and in all cases each region was sequenced on M13 subclones constructed from at least two independently isolated cDNA clones (Fig. 1). Sequence data was analyzed with Microgene sequencing analysis program (Beckman Instruments, Inc., Palo Alto, CA).

RNA Analysis. RNA isolated from human liver biopsy specimen from a normal liver, taken as part of a staging laparotomy for Hodgkin's disease, represents normal liver RNA. RNA was also isolated from a fresh postmortem liver of a victim who had been involved in a severe trauma. This RNA, a gift from Dr. G. Goldberger (Harvard Medical School, Boston, MA), was found to be greatly enriched for acute-phase reactants (26). These RNA samples are shown in Fig. 4. The samples in Fig. 5, A and B, were a generous gift from Dr. S. Karathanasis (Harvard Medical School, Boston, MA). The RNA was isolated from a human hepatoma cell line HepG2, postmortem specimens of liver, stomach, ileum, large

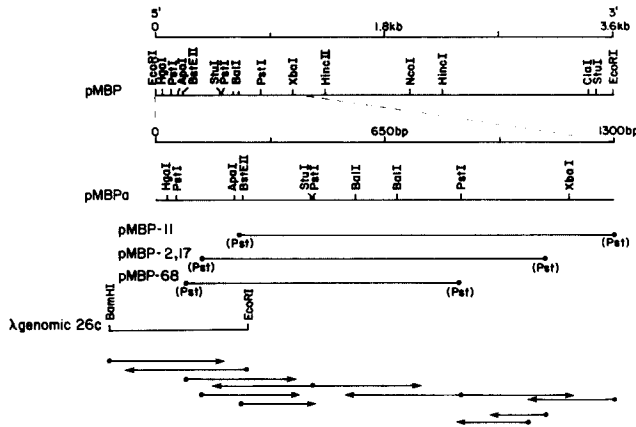


FIGURE 1. Restriction map of full-length human MBP-C cDNA is depicted as pMBP; pMBPa corresponds to the entire coding region and includes a putative polyadenylation tail; pMBP-2, 11, 17, and 68 represent independently isolated overlapping cDNA clones that were obtained from a GC-tailed library (30) cloned into a Pst I site. The arrows indicate the direction and extent of sequence analysis by the dideoxynucleotide chain termination method.

intestine, and spleen from a victim of a motor vehicle accident. In addition, fetal thymus tissue and 17- and 24-wk-old fetal liver tissue were also sources of RNA. The RNA samples were size fractionated on a formaldehyde agarose gel as described (30). RNA was transferred to nitrocellulose (30) and prehybridized in 50% formamide, 5× SSC, 5× Denhardt's solution, and 1% SDS at 42°C for 1 h. Radiolabeled human MBP cDNA or a Pst I genomic apolipoprotein clone that contains the entire coding region (31) (a gift from Dr. S. Karathanasis) was added at 10⁶ cpm/ml to hybridization solution containing 25 μg/ml of sonicated salmon sperm DNA. After overnight incubation at 42°C, the filter was washed at 0.2× SSC at 68°C. Slot blot analysis was also performed on 5 μg RNA isolated from HepG2 cells, the acute-phase liver, fetal liver, and normal liver, which are represented in lanes 1–4, respectively, in Fig. 5. The slot blot was hybridized with radiolabeled MBP cDNA and a radiolabeled partial cDNA that encodes part of the human third complement C3 (a gift from Dr. M. Carroll, Harvard Medical School, Boston, MA; reference 32), a known acute-phase reactant (33).

Results

Primary Structure of Human MBP

A radiolabeled rat cDNA (24) corresponding to the putative carbohydrate-binding region of the rat MBP-C was used to screen a human liver plasmid cDNA library. Five positive clones were isolated. Restriction map analysis and cDNA sequencing of clones pMBP 2, 11, 17, and 68 revealed that they contained identical overlapping sequence (Fig. 1). Comparison of the nucleotide sequence and derived amino acid sequence of rat with rat MBP-C (discussed in detail below) revealed that these clones did not contain the entire coding information for the human MBP. A Pst I–Bst E2 fragment (Fig. 1) that represented the most 5'-available sequence was radiolabeled and used to screen the human liver plasmid cDNA library but no clones containing additional 5' sequence were isolated. In parallel, a 3.5-kb cDNA clone, pMBP, was isolated from a HepG2 λ cDNA library (27). This clone spanned the shorter clones isolated from the liver library and contained the NH₂ terminus of the protein (Fig. 2, *boxed*). This was preceded by a typical hydrophobic signal peptide (discussed below) but did not include an initiator methionine. The 5' extent of the cDNA was confirmed by a genomic clone isolated from an EMBL 3A genomic library. Restriction mapping and Southern blot analysis revealed a genomic fragment that hybridized strongly

```

gct cgg taa ata tgt ttc att aac tga gat taa cct tcc ctg aqt
ttt ctc aca cca agg gag acc ATG TCC TGT TTC ATC ACT CCC TCT
Met Ser Cys Phe Ile Thr Pro Ser
CTT CTC CTG AGT ATG GTG GCA GCG TCT TAC TCA GAA ACT GTG ACC
Leu Leu Leu Ser Met Val Ala Ala Ser Tyr Ser Glu Thr Val Thr
TGT GAG GAT GCC CAA AAG ACC TGC CCT GCA GTG ATT GCC TGT AGC
Cys Glu Asp Ala Gln Lys Thr Cys Pro Ala Val Ile Ala Cys Ser
TCT CCA GGC ATC AAC GGC TTC CCA GGC AAA GAT GGG CGT GAT GGC
Ser Pro Gly Ile Asn Gly Phe Pro Gly Lys Asp Gly Arg Asp Gly
ACC AAG GGT AGA AAA GGG GGA ACA GGT CAA GGG CTC AGA GGC TTA
Thr Lys Gly Arg Lys Gly Gly Thr Gly Gln Gly Leu Arg Gly Leu
CAG GGC CCC CCT GGA AAG TTG GGG CCT CCA GGA AAT CCA GGG CCT
Gln Gly Pro Pro Gly Lys Leu Gly Pro Pro Gly Asn Pro Gly Pro
TCT GGG TCA CCA GGA CCA AAG GGC CAA AAA GGA GAC CCT GGA AAA
Ser Gly Ser Pro Gly Lys Gly Gln Lys Gly Asp Pro Gly Lys
AGT CCG GAT GGT GAT AGT AGC CCA GGC TGC CTC AGA AAG AAA AGC
Ser Pro Asp Gly Asp Ser Ser Pro Gly Cys Leu Arg Lys Lys Ser
TCT GCA AAC AGA AAT GGC ACG TAT CAA AAG TGC CTG ACC TTC TCT
Ser Ala Asn Arg Asn Gly Thr Tyr Gln Lys Cys Leu Thr Phe Ser
CTG GGC AAA CAA GTT GGG AAC AAG TTC TTC CTG ACC AAT GGT GAA
Leu Gly Lys Gln Val Gly Asn Lys Phe Phe Leu Thr Asn Gly Glu
ATA ATG ACC TTT GAA AAA GTG AAG GCC TTG TGT GTC AAG TTC CAG
Ile Met Thr Phe Glu Lys Val Lys Ala Leu Cys Val Lys Phe Gln
CCT CTG TGG CCA CCC CCA GGA ATG GCT GCA GAG AAT GGA GCC ATT
Pro Leu Trp Pro Pro Pro Gly Met Ala Ala Glu Asn Gly Ala Ile
CAG AAT CTC ATC AAG GAG GAA GCC TTC CTG GGC ATG CCT GAT GAG
Gln Asn Leu Ile Lys Glu Glu Ala Phe Leu Gly Met Asp Asp Glu
AAG ACA GAA GGG CAG TTT GTG GAT CTG ACA GGA AAT AGA CTG ACC
Lys Thr Glu Gly Gln Phe Val Asp Leu Thr Gly Asn Arg Leu Thr
TAC ACA AAC TGG AAC GAG GGT GAA CCC AAC AAT GCT GGT TCT GAT
Tyr Thr Asn Trp Asn Glu Gly Glu Pro Asn Asn Ala Gly Ser Asp
GAA CAT TGT GTA TTG CTA CTG AAA AAT GGC CAG TGG AAT GAC TCC
Glu His Cys Val Leu Leu Leu Lys Asn Gly Gln Trp Asn Asp Ser
CCT TGC TTC CAC CTC CCA TCT GCC GTC TGT GAG TTC CCT ATC TGA
Pro Cys Phe His Leu Pro Ser Ala Val Cys Glu Phe Pro Ile
agg gtc tgt gag ttc cct atc tga agg gtc ata tca ctc agg ccc
tcc ttg tct ttt tac tct acc aca gcc cac gta tgc ttg aaa gat
aat gat aat ata tca ttc tca gat cag tac ctg cag atg aaa ata
tca gat cag tac ctg cag atg aag ata aga cgg cat tta ttt ttc
cat tta caa caa aca cct gtg tgt tga gcc tta ctt tct gtt tgg
gta gag ggc tcc cct aat gac atg acc aca gtt taa tac cac agc
ttt ttt acc aag ttt cag gta tta aga aaa tct att ttg taa ctt
tct cta tga act ctg ttt tct ttc taa tga gac att aaa cca tgt
aaa aaa aaa aaa aaa aaa

```

FIGURE 2. Nucleotide sequence of 993 bases of human MBP-C cDNA. The typical hydrophobic signal sequence is underlined and followed by the putative NH₂ terminus of the protein which is boxed. The collagen region is depicted between the arrows. The termination codon which is boxed is followed by 183 bp of 3' untranslated sequence including a typical polyadenylation signal (43) followed by a poly(A) tail. The rest of the 3' untranslated region that makes up the full length cDNA is not shown.

to the 5' radiolabeled cDNA Pst 1-Bst E2. The sequence overlapped the available 5' cDNA sequence but extended further in the 5' direction. Examination of this sequence revealed an in-frame ATG followed by the signal sequence. The presence of an in-frame stop codon at nucleotide -36 upstream from the initiator methionine residue eliminates the possibility that the signal sequence is preceded by further extension.

Common Structural Features of Human MBP-C, Rat MBP-C, and Rat MBP-A

The sequence of the human MBP is strikingly homologous to the rat MBPs, in particular, rat MBP-C. As shown in Fig. 3, the human MBP and the rat MBP-C can be aligned with only three gaps. In this alignment the sequences of these two mature proteins are 51% homologous. An alignment between human MBP

1038 CHARACTERIZATION OF HUMAN MANNOSE-BINDING PROTEIN

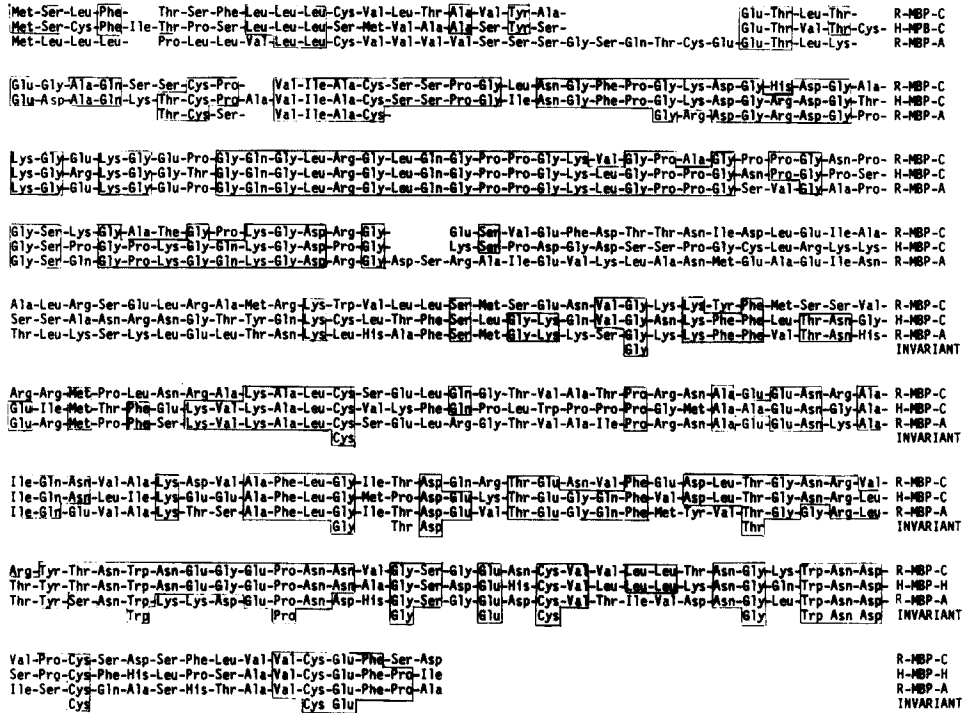


FIGURE 3. Comparison of animal lectin sequences. The sequences of the two rat MBPs, R-MBP-A and C, are compared with the human MBP-C H-MBP-C. The invariant residues represent sequences invariant in twelve lectin-like proteins including the human and rat hepatic asialoglycoprotein receptors (16), the avian hepatic receptor (16), the apoprotein of dog (18) and human surfactant (20); the NH₂ portion of a galactose-specific lectin isolated from the hemolymph of *S. perigonia* (21); a lectin isolated from the coelomic fluid of a sea urchin, *A. crassispina* (22); a chicken cartilage core proteoglycan protein (17); and the IgE Fc receptor (44). Residues that are identical between the MBPs, human, and rat are boxed to emphasize homology.

and the rat MBP-A with seven gaps allows a homology of 48% between these two native proteins. The overall organization of these proteins appears identical. Human MBP-C consists of a NH₂-terminal signal sequence followed by a short segment (21 amino acid residues) rich in cysteine residues, then a collagenous domain of 56 residues, and finally a COOH-terminal noncollagenous domain of 148 amino acids.

Signal Sequences

Since MBPs are found in the circulation, the presence of a typical hydrophobic signal sequence is an expected feature of these molecules. The human MBP-C signal sequence (underlined in Fig. 2) has the features of NH₂-terminal signal sequences found on almost all eukaryotic secretory proteins (34). The NH₂ terminal of the mature protein is the glutamic acid residue 20 (Fig. 2 boxed) based on the NH₂-terminal amino acid sequence of the MBP isolated from human serum (Baenziger, J., personal communication). This sequence is preceded by a serine that would conform to the general rule that the amino acid in position preceding cleavage by the signal peptidase tends to be a residue with a

small side chain-like serine or alanine (35). The lack of specific conservation of residues within the signal or at the cleavage boundary is similar to comparisons made between signal sequences in otherwise highly homologous proteins (35).

Interchain Disulfide Bonds

The presence of a number of cysteine residues in the short NH₂-terminal noncollagenous segment are found in both human and rat MBPs (Fig. 3). In the rat protein, these interchain disulfide bonds are removed when MBPs are digested with collagenase (24). When MBP purified from human sera is analyzed by PAGE under nonreducing conditions, it occurs as multimers of a 30-kD basic subunit (15, 25). Strong reduction and alkylation are required to reduce this multimeric form to the basic subunit indicating that these cysteines in this NH₂-terminal domain of human MBP-C are most likely involved in disulfide bond formation.

Collagen-like Domains

This region contains the strongest conservation and there are regions of near identity between the human MBP-C and both rat MBPs. There is a single identical interruption in the Gly-x-y repeat structure; the sequence Gly-Gln-Gly is found (residues 297–303) in a highly conserved portion of both collagenous domains. Other portions of the collagen-like domain resemble the triple helix-forming segments of collagen in that a large number of the x and y positions are occupied by prolines. Analysis of the human protein has not been undertaken, although in the rat, a high proportion of the prolines in the y position are hydroxyproline, therefore, one would predict similar hydroxylation of the human protein. The sequence Arg-Gly-Asp-Ser (RGDS) which is recognized by a family of cell surface receptors known as Integrins (reviewed in 36) is not found in the human MBP-C or the rat MBP-C, although it is present in the rat MBP-A.

Comparison of Human MBP with other Mammalian Carbohydrate Binding Proteins

Several recent reports have drawn attention to striking homologies between mammalian lectin-like proteins (16, 24). The primary structures of several animal lectins have been reported (17, 19, 21, 22, 24). Fig. 3 illustrates the invariant residues of 12 proteins that bear homology with the human MBP (these are listed in the Fig. 3 legend). The homologies are greatest in the COOH-termini of the mammalian lectin-like proteins that are the putative carbohydrate-binding domains. Although strong sequence homology exists, each protein has a distinct pattern of carbohydrate-binding specificity; the MBP's recognize mannose, the rat and human hepatic lectins are specific for galactose and *N*-acetylglucosamine, the chicken hepatic lectin recognizes *N*-acetylglucosamine. The binding specificity of the sea urchin lectin is not known. It would appear, too, that the lymphocyte Fc receptor for Ig E (IgE FcR) is a member of this lectin family and may also have lectin-like properties that may or may not relate to IgE binding. The cartilage proteoglycans and the apoproteins, SP 28–36, of pulmonary surfactant are most homologous to the galactose-binding proteins, and based on these predictions, these proteins do in fact interact with galactose (37).

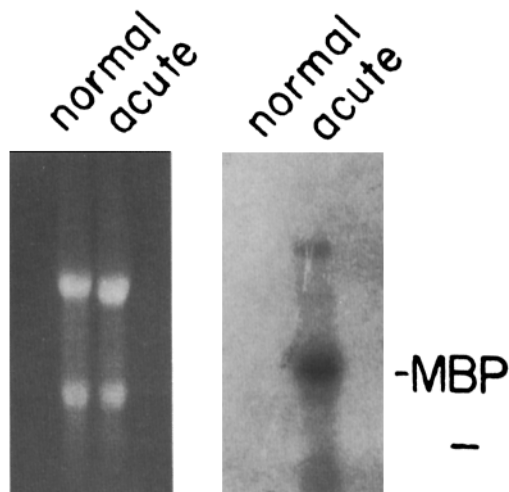


FIGURE 4. Northern blot analysis of RNA isolated from normal human liver (*normal*); and RNA isolated from liver that was enriched for acute-phase transcripts (*acute*). (*Left*) An ethidium bromide-stained gel before transfer onto nitrocellulose. 25 μ g of total RNA was loaded in both lanes and prominent 28S and 18S ribosomal RNA are seen. (*Right*) specific hybridization of human MBP cDNA that reveals predominant mRNA at 3.5 kb and lesser species of \sim 5.5 kb. Autoradiograph was exposed for 4 d.

Regulation. Variability of MBP levels isolated from different sera (from 100 ng/ml to 50 μ g/ml as determined by RIA; Ezekowitz, R. A. B., unpublished results) led us to investigate whether the hepatic synthesis of MBP may be regulated. Fig. 4 shows a Northern analysis of RNA isolated from normal liver after a liver biopsy for a staging laparotomy for Hodgkin's disease and RNA isolated from a patient who had 48 h before death suffered major trauma. The latter RNA was greatly enriched for acute-phase reactants (26). Radiolabeled human MBP cDNA hybridized only to the acute-phase RNA. A major species of RNA of \sim 3.5 kb represents the human MBP-C. A larger mRNA of \sim 5.5 kb is a consistent finding. This larger species may represent (*a*) additional 5' or 3' untranslated sequence; (*b*) a larger mRNA that codes for an homologous mannose-binding protein; or (*c*) a splicing intermediate.

Tissue Distribution. We next investigated whether MBP expression was restricted to liver. RNA was prepared from several different human tissues and cell lines. No specific hybridization of a radiolabeled MBP cDNA was observed to normal liver RNA (Fig. 5A, lane 4) after a 12-h exposure of the autoradiograph that contrasts to readily abundant apolipoprotein A1 transcripts (Fig. 5B, lane 4). Two mRNA species, 5.5 kb and 3.5 kb, were detected in acute-phase adult liver (lane 2; this RNA was isolated from a different liver than that shown in Fig. 4, and the autoradiograph was exposed overnight as compared with the 4-d exposure shown in Fig. 4). These differences are more clearly shown in the 36-h exposure of the slot blot (Fig. 5, *inset*). The pattern of MBP hybridization parallels the increase in the third complement component mRNA levels (Fig. 5, *inset, upper panel*). Both C3 mRNA and serum levels have been previously shown to be increased as part of the acute-phase response (33), although unlike MBP, base-line levels in normal liver and HepG2 cells are fairly abundant. No specific MBP mRNA transcripts were observed in RNA isolated from several other human tissues including small and large intestine, spleen, thymus, fetal liver, and stomach. However, apolipoprotein A1 mRNA was present in fetal and adult

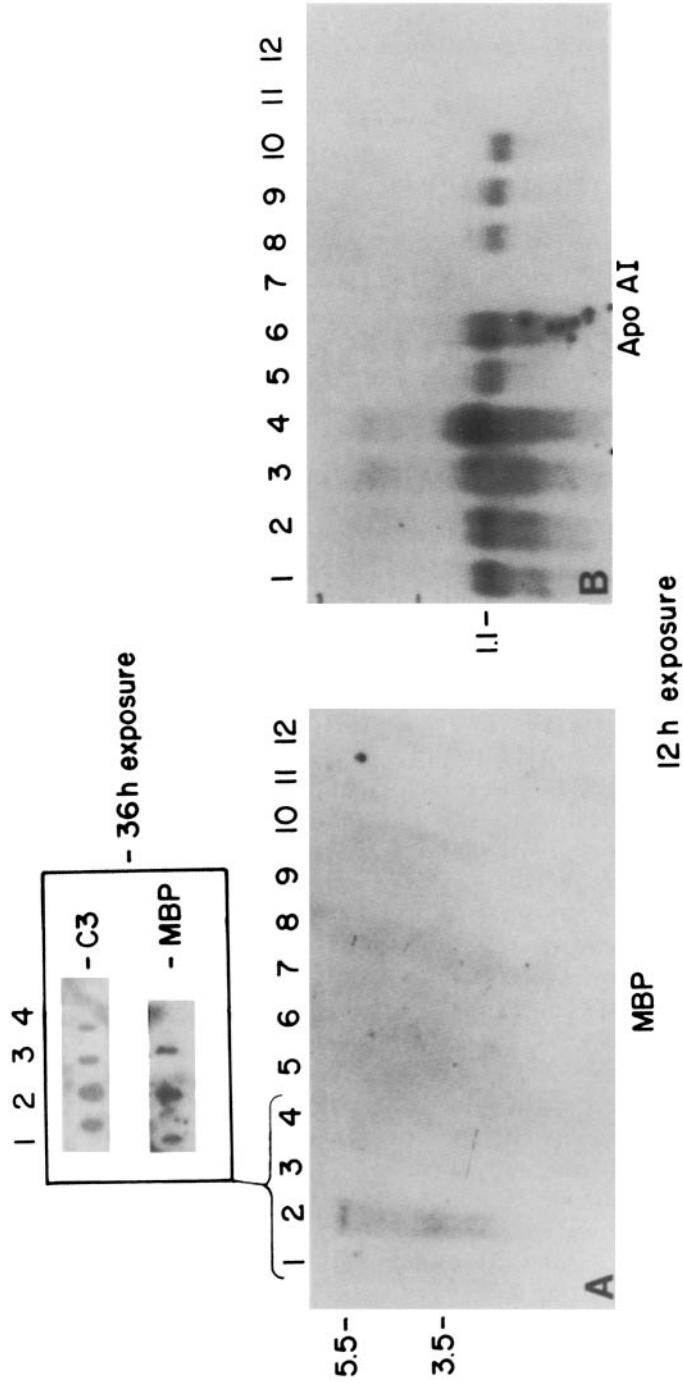


FIGURE 5. (A) Northern blot analysis of RNA isolated from various human tissues hybridized with radiolabeled MBP cDNA. Lane 1, HepG2; lane 2, adult acute phase liver; lane 3, adult normal liver; lane 4, 20-wk fetal liver; lane 5, 17-wk fetal liver; lane 6, 24-wk fetal liver; lane 7, stomach; lane 8, duodenum; lane 9, ileum; lane 10, large intestine; lane 11, thymus; lane 12, spleen; 15 μ g of total RNA was loaded in each lane. Autoradiograph was exposed overnight. (Inset) Slot blot analysis of 5 μ g of RNA that is represented in lanes 1-4. (Bottom) Hybridized with MBP; (top) hybridized with the third complement component, C3, a known acute-phase reactant (33). Autoradiograph was exposed for 36 h. (B) The identical Northern blot as described above hybridized with a radiolabeled Pst 1 fragment of the human apolipoprotein A1 gene containing the entire coding system (31).

liver, small and large intestine (31), and not in stomach, spleen, and thymus as expected (Fig. 5B). Human hepatoma cell line, HepG2, did express low levels of MBP RNA (observed after a 4-d exposure of autoradiograph; not shown); however, no specific hybridization was observed to human monocytes and human monocyte-derived macrophage RNA under conditions of high stringency (not shown).

Discussion

The human mannose-binding protein MBP-C can be divided into three domains: (a) an NH₂-terminal cysteine-rich domain that mediates interchain disulfide bonds; (b) a collagen-like domain; and (c) a COOH-terminal putative carbohydrate-binding domain. The derived amino acid sequence shows no canonical *N*-linked glycosylation site (38), therefore small amounts of carbohydrate if present must be attached by *O*-glycosidic linkage. Analysis of the NH₂-terminal amino acid sequence of the MBP isolated from human serum revealed that the sequence is Glu-Thr-Val-Thr-X-Glu-Asp-Ala-Gln-Lys (Baenziger, J., personal communication). This sequence is identical to the predicted NH₂-terminal amino acid sequence encoded by the MBP cDNA and therefore establishes MBP-C as the major mannose-binding protein in human serum. The fact that human MBP-C possesses a typical signal peptide, a feature of secreted proteins, and the fact that the hepatic synthesis of human MBP-C is upregulated as part of the acute-phase response, both support the contention that human MBP-C is synthesized in the liver and is secreted into the circulation.

The molecule is highly homologous to two rat MBPs, both in actual homology and organization of the molecule. The structure of the human MBP bears closest homology to the rat MBP-C, 51% overall homology with three gaps, and hence was designated human MBP-C. The human MBP-C bears 48% overall homology with the rat MBP-A (Fig. 3). Although human MBP-C shares a closer amino acid sequence homology with rat MBP-C, the human protein is functionally similar to the rat MBP-A. Like rat MBP-A (25), human MBP-C is predominantly a serum protein. Further, MBP isolated from rat serum (rat MBP-A) and human serum (human MBP-C), but not MBP isolated from rat liver (rat MBP-C), is able to activate complement by the classical pathway (25). It therefore appears that human MBP-C, while more homologous to rat MBP-C, is functionally equivalent to rat MBP-A.

The overall organization of collagen- and ligand-binding domains found in both human and rat MBP is also found in other proteins. These proteins include the asymmetric forms of acetylcholinesterase (39), complement protein Clq (a, b, c, chains; reference 40), and the apoprotein of human (20) and dog pulmonary surfactant (18). Clq and the surfactant apoproteins bear striking similarities with the MBPs. Pulmonary surfactant apoprotein has 30% overall homology with human MBP; however, to achieve this homology, seven gaps are required. The similarity between human MBP and complement Clq lies at the level of organization. Both of these proteins have short NH₂-terminal domains that are involved in interchain disulfide bond formation followed by triple helix-forming domains. Of particular interest from an evolutionary standpoint is the fact that the human MBP, rat MBP, apoprotein of pulmonary surfactant, and A and C chains of Clq

all have an aberration in the regular Gly-x-y-Gly repeats. It is known that the interruption is the site of an intron in human MBP (Herman, G., L. Day, E. Deignam, G. Bruns, and R. A. B. Ezekowitz, manuscript in preparation), in rat MBP (42), and in the human surfactant apoprotein gene (20). The collagen sequence is more closely related to that found in nonfibrillar collagen molecules than fibrillar collagen, in that distortions in the triple-helical segments are found in nonfibrillar collagen (39). These differences are exemplified at the gene level by the variability of exon size and intron site (Herman, G., L. Day, E. Deignam, et al., manuscript in preparation). This structure contrasts to the regular helical structure and exon size found in the structural fibrillar collagens (reviewed in 39).

The synthesis and secretion of the human MBP-C in the liver appears to be regulated rather than constitutive. MBP mRNA transcripts are barely detectable in normal liver samples but induction is observed in RNA isolated from independent human livers that have been exposed to acute stress (Figs. 4 and 5, respectively). This induction of specific mRNA correlates with variations in MBP serum levels as detected by RIA (Ezekowitz, R. A. B., unpublished observation). Purification of MBP-C from these sera also revealed the presence of high levels of C-reactive protein and serum amyloid A component (Baenziger, J., personal communication; and Ezekowitz, R. A. B., unpublished observation). The hepatic synthesis of these two acute-phase proteins is greatly induced in response to stress or inflammatory stimuli as reflected by the induction of mRNA and increased serum levels (reviewed in reference 23). The pattern of human MBP-C synthesis parallels that of these two well-characterized acute-phase reactants. This observation supports the inclusion of MBP-C as an acute-phase protein.

The precise function of the human MBP-C is not known. The human MBP binds mannose-rich yeasts (15) as well as the lipomannans of mycobacterium tuberculosis with high affinity (Ezekowitz, R. A. B., and P. Brennan, unpublished results) and is able to inhibit *in vitro* infection of susceptible cells by the human immunodeficiency virus (Ezekowitz, R. A. B., G. Groopman, and R. Byrn, manuscript submitted). These functions together with the circumstances under which MBP synthesis is induced as part of the acute-phase response, suggest a role for MBP in natural immunity. If the primary role of this protein is engagement of mannose-rich pathogens in the circulation, it would be expected that MBP-C may be a primitive opsonin able to distinguish mannose on pathogens from cell surface glycoproteins. Preliminary studies in progress appear to support this premise.

Summary

Mannose-binding proteins have been isolated from the liver of rats and humans and subsequently been found in the serum of rats, rabbits, and humans. We report the isolation of cDNA clones isolated from a human liver cDNA library that encodes a human mannose-binding protein. The primary structure has three domains: (a) an NH₂-terminal cysteine-rich segment of 19 amino acids which appears to be involved in the formation of interchain disulfide bonds that would stabilize multimeric forms of the protein; (b) a collagen-like region consisting of 19 repeats of the sequence Gly-x-y; and (c) a COOH-terminal putative carbohy-

drate-binding domain consisting of 148 residues. This human mannose-binding protein bears 51% overall homology (allowing three gaps) with a rat mannose-binding protein C and 48% homology (allowing seven gaps) with a rat mannose-binding protein A. Like these homologous rat proteins, the human mannose-binding protein COOH-terminal sequences are homologous to the carbohydrate recognition portion of several other lectin-like proteins including mammalian hepatic receptors, an insect-soluble hemolymph, and a sea urchin lectin found in coelomic fluid. The apoproteins of dog and human surfactant and the human lymphocyte IgE Fc receptor have not been shown to have lectin-like properties, yet by homology are members of this family of lectin-like proteins. The human mannose-binding protein is preceded by a typical hydrophobic signal sequence and its hepatic secretion is induced as part of the acute-phase response consistent with its probable role in host defense.

We wish to thank K. Drickamer for the generous gift of the rat MBP cDNA clones and for useful discussion and critical review of the manuscript. We are also grateful to S. Karathanasis for supplying human RNAs and the apolipoprotein probe; D. Kwiatkowski for the HepG2 cDNA library; and J. Baenziger for communicating unpublished results. S. Orkin's advice and support are also gratefully acknowledged.

Received for publication 17 July 1987 and in revised form 1 December 1987.

References

1. Stahl, R. D., and P. Schlessinger. 1980. Receptor-mediated pinocytosis of mannose/N-acetylglucosamine-terminated glycoproteins and lysosomal enzymes by macrophages. *Trends Biochem. Sci.* 5:194.
2. Stahl, R. D., T. Wileman, and V. L. Shepherd. 1985. The mannose receptor of macrophages: a current perspective. *In* Mononuclear Phagocytes, Characteristics, Physiology and Function. R. van Furth and Martinus Nijhoff, editors. 59-65.
3. Lennartz, M. R., T. Wileman, and P. D. Stahl. 1987. Isolation and characterization of a mannose-specific endocytosis receptor from rabbit alveolar macrophages. *Biochem. J.* In press.
4. Ezekowitz, R. A. B. 1988. The Macrophage. Biological Aspects in Natural Immunity. D. Nelson, editor. Academic Press, New York. In press.
5. Ezekowitz, R. A. B., R. Sim, M. Hill, and S. Gordon. 1984. Local opsonization by secreted macrophage components. Role of receptors for complement in uptake of zymosan. *J. Exp. Med.* 159:244.
6. Blackwell, J. M., R. A. B. Ezekowitz, M. B. Roberts, J. Y. Channon, R. B. Sim, and S. Gordon. 1985. Macrophage complement and lectin-like receptors bind *Leishmania* in the absence of serum. *J. Exp. Med.* 162:324.
7. Kawasaki, T., and G. Ashwell. 1977. Isolation and characterization of an avian hepatic binding protein specific for N-acetylglucosamine-terminated glycoproteins. *J. Biol. Chem.* 252:6536.
8. Mizuno, Y., Y. Kozutsumi, T. Kawasaki, and I. Yamashina. 1981. Isolation and characterization of a mannose-binding protein from rat liver. *J. Biol. Chem.* 256:4247.
9. Maynard, Y., and J. U. Baenziger. 1982. Characterization of a mannose and N-acetylglucosamine-specific lectin present in rat hepatocytes. *J. Biol. Chem.* 257:3708.
10. Colley, K. J., and J. Baenziger. 1987. Biosynthesis and secretion of rat core-specific lectin. *J. Biol. Chem.* 262:3415.
11. Summerfield, J. A., and M. E. Taylor. 1986. Mannose-binding proteins in human

- serum: identification of mannose-specific immunoglobulins and a calcium dependent lectin, of broader carbohydrate specificity secreted by hepatocytes. *Biochim. Biophys. Acta.* 883:197.
12. Wild, J., D. Robinson, and B. Winchester. 1983. Isolation of mannose-binding proteins from human liver. *Biochem. J.* 210:167.
 13. Kawasaki, T., R. Eton, and I. Yamashura. 1978. Isolation and characterization of a mannose binding protein from rabbit liver. *Biochem. Biophys. Res. Commun.* 81:1018.
 14. Townsend, R., and P. Stahl. 1981. Isolation and characterization of a mannose/*N*-acetylglucosamine/fucose binding protein from rat liver. *Biochem. J.* 194:209.
 15. Kawasaki, N., T. Kawasaki, and I. Yamashura. 1983. Isolation and characterization of a mannose binding protein from human serum. *J. Biochem.* 94:937.
 16. Drickamer, K. 1987. Structure and biosynthesis of membrane receptors which mediate endocytosis of glycoproteins. *Kidney Int.* 32 (Suppl. 23):67.
 17. Shigaku, S., T. Takaaki, R. A. Kosher, and M. L. Tanzier. 1986. Cloning and sequence analysis of a partial cDNA for chicken cartilage proteoglycan core protein. *Proc. Natl. Acad. Sci. USA.* 83:5081.
 18. Benson, B., S. Hawgwood, J. Schilling, J. Clements, P. Danim, B. Cordell, and R. T. White. 1985. Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino acid sequence. *Proc. Natl. Acad. Sci. USA.* 82:6379.
 19. Martin, G. R., R. Timpl, P. K. Muller, and K. Kuhn. 1985. The genetically distinct collagens. *Trends Biochem. Sci.* July 285.
 20. White, R. T., D. Danom, J. Miller, K. Spratl, J. Schilling, S. Hawgwood, B. Benson, and B. Cordell. 1985. Isolation and characterization of the human pulmonary surfactant apoprotein gene. *Nature (Lond.)*. 317:361.
 21. Takahashi, H., H. Komano, N. Kawaguchi, N. Kitamura, S. Nakanishi, and S. Naton. 1985. Cloning and sequencing of a cDNA of Sarcophaga peregrina humoral lectin induced by injury of the body wall. *J. Biol. Chem.* 260:12228.
 22. Giga, Y., I. Atsushi, and K. Takahashi. 1987. The complete amino acid sequence of echinoiden, a lectin from the coelomic fluid of the sea urchin *Anthocidaris crassispina*. *J. Biol. Chem.* 262:6197.
 23. Pepys, M. B., and M. L. Baltz. 1983. Acute-phase proteins with special reference to C-reactive protein and related proteins (Pentraxins). *Adv. Immunol.* 34:141.
 24. Drickamer, K., M. S. Dordal, and L. Reynolds. 1986. Mannose-binding proteins isolated from rat liver contain carbohydrate-recognition domains linked to collagenous tail. *J. Biol. Chem.* 261:6578.
 25. Ikeda, K., T. Sannoh, N. Kawasaki, T. Kawasaki, and I. Yamashina. 1987. Serum lectin with known structure activates complement through the classical pathway. *J. Biol. Chem.* 262:7451.
 26. Woods, D. E., A. F. Markham, A. T. Ricker, G. Goldberger, and H. R. Colten. 1982. Isolation of cDNA clones for human complement protein factor B, a class III major histocompatibility complex gene product. *Proc. Natl. Acad. Sci. USA.* 79:5661.
 27. Kwiatkowski, D. J., T. P. Stossel, S. H. Orkin, J. E. Mole, H. R. Colten, and H. L. Yin. 1986. Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin-binding domain. *Nature (Wash. DC)*. 323:455.
 28. Sanger, F., A. S. Nicklen, and A. R. Coulson. 1977. DNA sequencing by chain termination. *Proc. Natl. Acad. Sci. USA.* 74:5463.
 29. Biggin, M. D., T. J. Gibbon, and G. F. Hong. 1983. Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA.* 80:3963.
 30. Maniatis, T., F. Fitch, and S. Sambrook. 1983. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.

31. Karathanasis, S. K. 1985. Apolipoprotein multigene family Tandem organization of human apolipoprotein A1, C3 and A4 genes. *Proc. Natl. Acad. Sci. USA.* 82:6374.
32. de Bruijn, M., and G. Fey. 1985. Human complement component C3: *Proc. Natl. Acad. Sci. USA.* 82:708.
33. Darlington, G. J., D. R. Wilson, and L. B. Lachman. 1986. Monocyte-conditioned medium, interleukin 1, and tumor necrosis factor stimulate the acute-phase response in human hepatoma cells in vitro. *J. Cell Biol.* 103:787.
34. von Heijne, G. 1982. Signal sequences are not uniformly hydrophobic. *J. Mol. Biol.* 159:537.
35. von Heijne, G. 1983. Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* 133:17.
36. Hynes, R. D. 1987. Integrins. A family of cell surface receptors. *Cell.* 48:549.
37. Haagsman, H. P., S. Hawgwood, T. Sargeant, D. Buckley, R. T. White, K. Drickamer, and B. J. Benson. 1987. The major lung surface protein, SP28-36 is a calcium-dependent carbohydrate-binding protein. *J. Biol. Chem.* 262:13877.
38. Kornfeld, R., and S. Kornfeld. 1985. Assembly of asparagine linked oligosaccharides. *Annu. Rev. Biochem.* 54:631.
39. Rossenberry, T. L., and J. M. Richardson. 1977. Structure of 185 and 145 acetylcholinesterase: identification of collagen-like subunits that are linked by disulphide bonds to catalytic subunits. *Biochemistry.* 16:3550.
40. Reid, K. B. M. 1983. Proteins involved in the activation and control of the two pathways of human complement. *Biochem. Soc. Trans.* 11:1.
41. Drickamer, K., and V. McCreary. 1987. Exon structure of a mannose-binding protein gene reflects its evolutionary relationship to the asialoglycoprotein receptor and nonfibrillar collagens. *J. Biol. Chem.* 262:2582.
42. Proudfoot, N., and G. Brownlee. 1976. 3' non-coding region sequences in eukaryotic mRNA. *Nature (Lond.).* 263:211.
43. Ikuta, K., M. Takami, C. W. Kim, T. Honjo, T. Miyoski, Y. Tagaya, T. Kawabe, and J. Yodou. 1987. Human lymphocyte Fc receptor for IgE. Sequence homology of its cloned cDNA with animal lectins. *Proc. Natl. Acad. Sci. USA.* 84:819.