Mannose-binding Proteins Isolated from Rat Liver Contain Carbohydrate-recognition Domains Linked to Collagenous Tails

COMPLETE PRIMARY STRUCTURES AND HOMOLOGY WITH PULMONARY SURFACTANT APORTEIN

(Received for publication, January 15, 1986)

Kurt Drickamer‡, Margaret S. Dordal, and Lisa Reynolds

From the Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637

Preparations of mannose-binding protein isolated from rat liver contain two distinct but homologous polypeptides. The complete primary structures of both of these polypeptides have been determined by sequencing of peptides derived from the proteins, isolation and sequencing of cDNAs for both proteins, and partial characterization of the gene for one of the proteins. Each polypeptide consists of three regions: (a) an NH₂-terminal segment of 18-19 amino acids which is rich in cysteine and appears to be involved in the formation of interchain disulfide bonds which stabilize dimeric and trimeric forms of the protein, (b) a collagen-like domain consisting of 18-20 repeats of the sequence Gly-X-Y and containing 4-hydroxyproline residues in several of the Y positions, and (c) a COOH-terminal carbohydrate-binding domain of 148-150 amino acids. The sequences of the COOH-terminal domains are highly homologous to the sequence of the COOH-terminal carbohydrate-recognition portion of the chicken liver receptor for N-acetylglucosamine-terminated glycoproteins and the rat liver asialoglycoprotein receptor. Each protein is preceded by a cleaved, NH₂-terminal signal sequence, consistent with the finding that this protein is found in serum as well as in the liver. The entire structure of the mannose-binding proteins is homologous to dog pulmonary surfactant apoprotein.

Proteins capable of recognizing specific oligosaccharides have been isolated from a number of different organs in a range of vertebrate species (1). These proteins fall into several categories. One group consists of integral membrane proteins which are involved in the intracellular routing of glycoproteins to specific destinations, particularly the lysosomes. Examples of these proteins include the mammalian hepatic receptor for asialoglycoproteins, the avian hepatic receptor for agalacto-glycoproteins, and the ubiquitous mannose-phosphate receptors (1, 2). A second group of small, water-soluble galactose-binding proteins have been extracted from many different tissues; although they are not integral membrane proteins, their localization within the cytoplasm or outside the cells remains a matter of controversy (3, 4). Finally, the serum amyloid protein, a precursor to the protein component of amyloid plaques and a member of the pentraxin group of serum proteins, specifically binds to galactose-containing polymers (5).

The primary structures of several of these carbohydrate-binding proteins are known. Five membrane-bound glycoprotein receptors, including two distinct asialoglycoprotein receptors isolated from rat liver and two found in human liver, as well as the chicken agalacto-glycoprotein receptor, are homologous to each other, particularly in a COOH-terminal domain of 130-150 amino acids which contains the carbohydrate-binding activity (6-8). Each of these proteins is anchored to the cell membrane by a membrane-spanning stretch of hydrophobic amino acids and an NH₂-terminal cytoplasmic tail (8). In contrast, the primary structures of mouse and human serum amyloid protein are highly homologous to each other and to the acute phase reactant referred to as C-reactive protein (9), but do not appear to be structurally related to the membrane receptors.

A mannose-binding protein (MBP¹) isolated from rabbit and rat liver was originally believed to be involved in uptake of mannose-containing glycoproteins by macrophages (9-12); however, the fact that this protein is found exclusively in the parenchymal cells of the liver rules out this function (13, 14). A MBP was subsequently isolated from the serum of rabbits and rats as well as humans (15, 16); the serum and liver MBPs appear to be identical based on binding specificity, pH and calcium ion dependence of binding, immunological cross-reactivity, and identical mobilities on SDS-polyacrylamide electrophoresis gels.

The MBPs share several important characteristics with the membrane-bound receptors found in hepatocytes, such as calcium ion dependence and loss of binding activity below pH 6.0. On the other hand, the MBPs are water-soluble and can be extracted from liver in the absence of detergent, which is in contrast to the membrane-bound receptors. The presence of MBP in the circulation suggested that this protein might be more related to serum amyloid protein.

Structural analysis of MBP isolated from rat liver was undertaken to determine the degree to which this protein is related to the previously characterized carbohydrate-binding proteins. The results of this analysis reveal that there are two rat MBPs which are homologous to each other and to the carbohydrate-binding portions of the membrane receptors. The remainder of each polypeptide consists largely of colla-

¹The abbreviations used are: MBP, mannose-binding protein; SNE, sodium dodecyl sulfate; PTH, phenylthiohydantoin.
gen-like sequences. Some possible functions for MBPs suggested by their structural organization are discussed.

EXPERIMENTAL PROCEDURES

RESULTS

Organization of MBP Polypeptides

Multiple Polypeptides in Preparations of MBP—MBP can be isolated by several different procedures. The central purification step is affinity chromatography either on an immobilized glycoprotein which contains a high proportion of mannosidase isolated by several different procedures. The central purification step is affinity chromatography either on an immobilized glycoprotein which contains a high proportion of mannose residues (such as yeast mannan or invertase) or on columns constructed directly with mannose coupled to agarose. In earlier experiments, MBP was extracted from an acetone powder of rat liver with Triton X-100 (9, 11). Subsequently, it was found that extraction with high salt alone releases MBP in good yield (10, 13). The material isolated by any of these procedures appears to be the same: in each case, a polypeptide of $M_r = 29,000$ is isolated in what appears to be homogeneous form. However, upon reduction and alkylation, the presence of two polypeptides with slightly different mobilities on SDS-gel electrophoresis is revealed (see Fig. 1). This figure also demonstrates that the species which migrates faster (designated MBP-A) is enriched in a fraction which precipitates upon dialysis against water, and is essentially absent from the fraction which remains soluble. This provides a useful means of partially resolving the two different species of MBP. The two forms of MBP have been found in preparations of MBP in which either invertase-Sepharose or mannose-Sepharose was employed as the affinity resin after starting with either detergent or high salt extraction. In all cases, the more slowly migrating species (MBP-C) is the predominant form.

Direct evidence for the presence of two distinct polypeptides in the MBP preparations was obtained by sequence analysis of the two fractions shown in Fig. 1. The results, shown in Table 1, demonstrate that a unique sequence is obtained for the pure MBP-C fraction, whereas a double sequence is obtained when some MBP-A is present with the MBP-C. Because the MBP-C sequence can be determined by itself, the sequence of MBP-A probably reflects the low abundance of the MBP-A form. The absence of a second sequence derived from either MBP-A or MBP-C. The yield of the fragment indicates that it must be derived from the more abundant MBP-C. The absence of a second sequence derived from MBP-A probably reflects the low abundance of the MBP-A fragment. The sequence of intact MBP-A could be discerned.

Presence of Collagen-like Regions within the MBPs-In the course of generating additional sequence information for MBP-A and C, a number of tryptic peptides were isolated by two-dimensional thin layer peptide mapping (Fig. 2) and analyzed by automated Edman degradation (Tables II and III). Two of these peptides (T2 and T7) consist of repeats of the sequence Gly-X-Y, which is characteristic of the sequences found in portions of collagen which form triple helices (33). In addition, several of the Y positions are occupied by 4-hydroxyproline residues, another characteristic of collagenous sequences (33). The amino acid composition of the MBP preparations reveals a high proportion of glycine (9), as would be expected if it contains a collagen-like sequence. When a special analysis was conducted for 4-hydroxyproline, it was found to be present at 1.3 mol \%.

In order to localize the putative collagen-like portion of MBP, bacterial collagenase was used to probe for collagen-like sequences. This enzyme preferentially cleaves polypeptides after proline in the sequence Gly-X-Pro-Gly (36). The results of this digestion are shown in Fig. 3. Most of the MBP polypeptides are cleaved by collagenase, leaving a fragment (C-1) of apparent molecular weight 20,000. The absence of an additional polypeptide fragment visible on the gels suggests that roughly one-third of each $M_r = 29,000$ polypeptide is degraded by collagenase. The C-1 fragment can be repurified by affinity chromatography on invertase-Sepharose (Fig. 3, lane 3), which indicates that the mannose-binding domain of the MBP resides in the collagenase-resistant portion of the molecule.

Fragment C-1 was subjected to sequence analysis with the results shown in Table III. The first 9 residues of the sequence consists of collagen-like Gly-X-Y repeats, following which the sequence diverges from this pattern. This sequence is completely unrelated to the sequence found at the NH$_2$ terminus of either MBP-A or MBP-C. The yield of the fragment indicates that it must be derived from the more abundant MBP-C. The absence of a second sequence derived from MBP-A probably reflects the low abundance of the MBP-A fragment. The sequence of intact MBP-A could be discerned.
only when this form of MBP was enriched by the differential precipitation procedure (Fig. 1 and Table I), and was not detected when total MBP was subjected to sequence analysis (data not shown). Thus a collagenase-resistant fragment of MBP-A would not be detectable in this experiment.

Taken together, these results suggest that the NH$_2$-terminal one-third of at least MBP-C consists of a triple helical region of protein, and that the carbohydrate-binding portion of the receptor lies in the COOH-terminal two-thirds of the molecule.

Interchain Disulfide Bonds—Evidence for the multimeric nature of MBP was obtained by analysis of the protein under nonreducing conditions. An SDS-polyacrylamide gel of an unreduced preparation of MBP (Fig. 3, lane 4) reveals the presence of higher molecular weight bands with apparent molecular weights of 57,000 and 78,000. These bands appear to represent disulfide bond-stabilized dimers and trimers of the MBP polypeptides. Similar ratios of dimers and trimers have been observed by others (14). Gel filtration analysis of the MBP preparation suggests that the native molecule (M, = 194,000) consists of six polypeptide subunits (9), which would be consistent with the presence of either two disulfide-linked trimers or three disulfide-linked dimers. Although it is possible that the trimers are formed of MBP-C subunits and the dimers of MBP-A subunits, the staining intensities do not appear to be consistent with this interpretation, since the relative abundance of dimers compared to trimers is greater than the relative abundance of MBP-A compared to MBP-C (Figs. 1 and 3).

The location of the cysteines involved in the disulfide bonds was determined by examining the mobility of the collagenase-resistant portion of MBP-C under nonreducing conditions (Fig. 3, lane 5). The mobility of this fragment is not significantly affected by reduction, indicating that the interchain disulfide bonds are not located within the carbohydrate-binding domain. This suggests that they are likely to be near the NH$_2$-terminal end of the polypeptide, a conclusion which is consistent with the sequence results presented below.

Primary Structures of Two MBPs

Isolation and Characterization of cDNA Clones—Oligonucleotide probes designed to be complementary to the MBP mRNA were used to screen rat liver cDNA libraries. In order to obtain a suitable sequence from which to design the probe, a contiguous stretch of noncollagenous amino acid sequence was desired. This was obtained by digesting the collagenase-resistant fragment with clostripain and isolating a large fragment (Cl-1) from the resulting mixture (Fig. 4). When this fragment was sequenced, 29 residues were identified (Table III). Two regions from which the mRNA sequence could be predicted with relatively little ambiguity were selected (Fig. 5). At positions of 2-fold ambiguity, both possible sequences were synthesized by including a mixture of both bases in the coupling reaction. At positions of 4-fold ambiguity, T was selected based on our previously employed reasoning that this might allow an A/T or G/T base pair with some stability, or at worst C/T or T/T pairs which would not be extremely destabilizing (27). In probe 2, a G was placed opposite the first base of the single arginine codon, since C is most commonly found in this position. This made possible the synthesis of 29 base oligonucleotides, thus taking advantage of the high signal to noise ratio which can be obtained using such relatively large probes (37).

When a rat liver cDNA library prepared in phage $\lambda$GT11 (20) was screened using oligonucleotide probe 1, two positive plaques were detected out of 200,000 plaques screened. Upon plaque purification, one of these was found to hybridize to probe 2 as well as probe 1 and was selected for further analysis. Sequencing revealed that this clone contains a partial copy of the MBP-C mRNA (see Fig. 6). A restriction fragment which includes the sequence encoding the COOH-terminal 120 amino acids of the protein was used to screen a rat liver cDNA library constructed in a plasmid vector. A single clone of MBP-C was obtained from this library out of 100,000 colonies screened. The complete sequence of the insert in this clone was established (Fig. 7), revealing that it includes the entire protein-coding portion of the MBP-C mRNA.

When the $\lambda$GT11 library was rescreened using the same restriction fragment probe, two categories of phage were detected. One group, which gave a stronger signal upon hybridization, consisted of partial cDNAs for MBP-C. A total of 15 plaques, representing six independent phage, were detected in 1,000,000 plaques screened. In addition, a group of three more weakly hybridizing plaques, representing two independent phage, was identified. The inserts from this group of phage were excised with EcoRI and ligated into plasmid pC646 for further analysis. In all cases, sequence analysis revealed that these phage contain coding information for MBP-A. The longest insert, in plasmid pMHL-A, was completely sequenced, with the results shown in Fig. 8.

Comparison of cDNA and Protein Sequences—The sequences of the proteolytic fragments of MBP discussed in the preceding section are shown along with the amino acid sequences deduced from the cDNA sequences in Figs. 7 and 8. All of the tryptic fragments obtained can be accounted for in one of the two predicted sequences. This includes a number of double sequences in which two tryptic fragments were not resolved by peptide mapping and were simultaneously subjected to Edman degradation. Peptide T7 is found in both MBP-A and MBP-C. In addition, the NH$_2$-terminal sequence of the collagenase-resistant portion of MHL-C (C-1) and the sequence of the large clostripain fragment used to generate the oligonucleotide probes (Cl-1) match exactly the sequences predicted from the cDNA. The calculated amino acid compositions of both MBPs are very similar to the composition determined for the mixture (10, 13). The calculated molecular weights are 24,000 for MBP-C and 23,500 for MBP-A. These values are somewhat lower than the values estimated from gel electrophoresis. Two factors may account for this discrepancy: the collagen-like sequences may lead to anomalous mobility on the gels, and there may be some carbohydrate attached to the mature proteins (see below).

Two differences between the protein and nucleic acid sequences were observed. As noted above, tryptic peptides T2 and T7 were found to contain several residues of hydroxyproline. These residues are encoded as proline residues in the cDNA. This is expected, since hydroxylation occurs as a post-translational modification (33). The amount of hydroxyproline in the MBP preparation suggests that additional proline residues in MBP-A and MBP-C are probably modified in this manner. Interestingly, 2 residues of peptide T2 which were not identified by protein sequencing were found to be lysine residues in the cDNA-derived sequence. It is possible that these residues may be hydroxylated, as is commonly the case in collagenous sequences (33). Although the carbohydrate composition of MBP has been found to be very low (13), the amounts detected leave open the possibility that glycosylated hydroxylsine residues might be present. There are no Asn-X-Ser(Thr) sequences in either MBP-A or C, so the small amount of carbohydrate present must be attached by an α-glycosidic linkage (38). Attempts to...
determine the amount of hydroxylysine present in the MBP preparation were unsuccessful.

Comparison of the NH2-terminal sequences of MBP-A and MBP-C with the sequences encoded by the cDNAs reveals that each protein is preceded by a leader peptide. In the case of MBP-C, this extension has the typical hydrophobic character of a signal sequence for directing the insertion of this protein through a microsomal membrane (39). The site of signal cleavage follows the general pattern of recognition seen in other signal peptides, since the residues at positions -1 and -3 (Ala and Val) are amino acids commonly found at these critical positions (40). The presence of an in-frame stop codon 20 nucleotides upstream from the indicated initiator methionine residue eliminates the possibility that the signal sequence is preceded by a further extension. It is noteworthy that this stop codon is immediately preceded by a methionine codon, indicating that in this case protein translation is not initiated at the first AUG sequence in the mRNA.

Characterization of MBP-A Signal Sequence by Sequencing of Genomic Clones—The longest available cDNA for MBP-A does not cover the entire protein-coding portion of the mRNA, so the NH2-terminal sequence of the initial translation product was determined by characterizing a portion of the gene for this protein. A rat genomic library was screened using a restriction fragment derived from the 5' end of plasmid pMBP-A (see Fig. 6). Several overlapping phage were isolated and characterized.4 The sequence of the relevant portion of the MBP-A gene is presented in Fig. 9. This sequence overlaps the 5' end of the cDNA sequence, and reveals that MBP-A is also preceded by a hydrophobic signal sequence. The absence of any potential splice acceptor sites (41) between the initiator methionine codon and the 5' end of the cDNA sequence insures that no intron interrupts this portion of the sequence, although a possible splice acceptor site is found 10 nucleotides upstream of the AUG codon.

DISCUSSION

Common Structural Features of MBP-A and MBP-C

The sequences of the two MBPs are striking homologous to each other throughout their length. As shown in Fig. 10, the two proteins can be aligned with only four gaps. In this alignment, the sequences of the mature proteins are 56% identical. As summarized in Fig. 11, the overall organization of the proteins is also identical. Each consists of an NH2-terminal signal sequence which is removed from the mature protein, followed by a short segment (18–20 amino acid residues) rich in cysteine residues, then a collagenous domain of 53–59 residues, and finally a COOH-terminal noncollagenous domain of 148–150 amino acids. As expected from the small difference in mobility of the reduced and carboxymethylated proteins, MBP-C is slightly longer than MBP-A.

Signal Sequences—Since the MBPs are found in the circulation and inside microsomes as well as possibly at the cell surface (13, 15), the presence of signal sequences is an expected feature of the molecules. The signal sequences of the two MBPs are the least conserved portions of the molecules; however, each has the general features of NH2-terminal signal sequences found on almost all eukaryotic secretory proteins (39, 42). In addition to being hydrophobic in character, both signal sequences contain cysteine residues near the midpoint. This is a common, although not universal, characteristic of signal sequences (42). The sites of cleavage by signal peptidase conform to the general rule that the amino acid in position -1 tends to be a residue with a small side chain (alanine and serine in MBP-C and MBP-A, respectively). The lack of specific conservation of residues within the signal or at the cleavage boundary is typical of comparisons made between signal sequences in otherwise highly homologous proteins (42).

Interchain Disulfide Bonds—The presence of a number of cysteine residues in the short NH2-terminal noncollagenous segment of both proteins is consistent with the fact that the interchain disulfide bonds are removed when the MBPs are digested with collagenase. Since both disulfide-linked dimers and trimers can be observed when the gels are run under nonreducing conditions, at least two of the cysteines in this NH2-terminal domain must be involved in disulfide bonds. The pattern of multimer formation may be somewhat heterogeneous, since both dimers and trimers are seen. As discussed above, the relative abundance of these species is not consistent with the hypothesis that one represents MBP-C and the other MBP-A.

Collagen-like Domains—Portions of the collagen-like domains of the two MBPs show extremely strong conservation. Of particular interest is the presence of a single, identical interruption in the Gly-X-Y repeat structure: the sequence Gly-Gln-Gly is found in a highly conserved portion of both collagenous domains. In all other portions of the collagen-like domains, the sequences resemble the triple helix-forming segments of collagen in that a large number of the X and Y positions are occupied by proline residues. In addition, the protein sequence work indicates that many of the proline residues in the Y positions are hydroxylated. The composition of 1.3 mol % 4-hydroxyproline indicates that approximately 3 residues are present per polypeptide chain. Since there are four prolines in Y positions in MBP-A and five in MBP-C, a significant fraction of the potentially hydroxylated residues must actually be derivatized. In both MBP-A and C, one of the prolines in a Y position directly precedes the irregularity in the Gly-X-Y repeat discussed above; this may prevent its hydroxylation.

Role of Multiple MBPs—The existence of two distinct proteins in preparations of MBP raises the question of whether these proteins have distinct biological functions. For example, it is possible that the serum form of MBP may be enriched in MBP-A compared with MBP-C; in this case, the MBP-A isolated from liver might represent "contamination" from circulating protein. Alternatively, the two MBPs may have distinguishable binding specificities. The availability of clones for the two proteins will make it possible to investigate this possibility using the methods of in vitro transcription and translation (43).

Comparison of MBPs with Other Mammalian Carbohydrate-binding Proteins

Identification of a Common Carbohydrate-recognition Domain—The collagenase digestion data demonstrate that the mannose-binding portion of MBP-C is located in the COOH-terminal collagenase-resistant domain. Because of the homologous organization of the MBPs, the same is probably true of MBP-A as well. The carbohydrate-binding domains of the rat and chicken endocytic receptors for galactose- and N-acetylglucosamine-terminated glycoproteins have also been localized to the COOH-terminal portions of these molecules (8, 43). It was, therefore, of interest to compare the sequences of these carbohydrate-recognition domains. As shown in Fig. 10, the sequences are strikingly homologous with each other. The COOH-terminal portion of the rat asialoglycoprotein receptor

4 A complete characterization of the gene encoding MBP-A will be published at a later date (V. M. McCready and K. Drickamer, manuscript in preparation).
**Fig. 10. Comparison of animal lectin sequences.** Sequences of the two MBPs are compared with a COOH-terminal fragment of the rat liver asialoglycoprotein receptor produced by clostripain digestion (rat hepatic lectin, RHL-1 (5, 43)), the COOH-terminal fragment of the chicken hepatic receptor for N-acetylglucosamine-terminated glycoproteins produced by subtilisin digestion (chicken hepatic lectin, CHL (8, 19)), the apoprotein of dog pulmonary surfactant (DPS (45)), and the NH2-terminal portion of a galactose-specific lectin isolated from the hemolymph of Sarcophaga peregrina (invertebrate soluble lectin, ISL (49)). Residues which can be aligned with either MBP by introducing five gaps to give 23% sequence identity, whereas the chicken receptor can be aligned with four gaps to show 32% identity. There are certain residues which are conserved in all of the carbohydrate-binding domains, particularly around some of the cysteine residues. These residues may be involved in disulfide bond formation. These residues are also conserved in other membrane-bound glycoprotein receptors, such as the minor form of the rat asialoglycoprotein receptor (6) and the two forms of human liver asialoglycoprotein receptor (7). Overall, 19 residues in the carbohydrate-recognition domain are conserved in all these proteins.

As summarized in Fig. 11, the COOH-terminal approximately 130 amino acid residues of these proteins represent a carbohydrate-recognition domain which has been combined, in the course of evolution, with several other different types of domains: the carbohydrate-recognition domain is attached to a collagenous domain in the MBPs, and to a membrane anchor in the case of the endocytic receptors. This represents a striking example of how conserved domains can be combined in the construction of proteins which serve distinct functions but which also share at least one characteristic activity.

In spite of the strong sequence conservation in the carbohydrate-recognition domain, there are certain residues which are conserved in all of the carbohydrate-binding domains, particularly around some of the cysteine residues. These residues may be involved in disulfide bond formation.

---

6 D. Farrell, E. Hsueh, and K. Drickamer, unpublished observations.
as homologous carbohydrate-recognition domains lie at the COOH termini of all of the proteins except the lectin from S. peregrina. Collagenous domains which are likely to form triple helices are shown as open boxes. Membrane-spanning segments which also serve as internal signal sequences are denoted by thickened lines. NH$_2$-terminal, cleavable signal sequences are shown as wavy lines. Membrane-spanning segments which also serve as internal signal sequences are denoted by thickened lines. Abbreviations are the same as in Fig. 10.

Carbohydrate-recognition domains, it should be noted that each of these proteins has a distinct pattern of carbohydrate-binding specificity: the MBPs bind mannose and N-acetylgalactosamine (10), whereas the rat asialglycoprotein receptor is specific for galactose and N-acetylgalactosamine, and the chicken endocytic receptor recognizes N-acetylgalactosamine (1), it will require further investigation to determine which residues in the carbohydrate-recognition domain are responsible for conferring these distinct specificities on the binding domains.

Multimeric Structure—The presence of triple helices and disulfide-linked trimers indicates that the MBPs must be at least trimers in structure. Sedimentation analysis yields a molecular weight of 200,000, which suggests that two trimers may be associated in a hexamer in the native molecule. The exact nature of the disulfide bonding pattern and the question of whether the two MBPs form independent or mixed hexamers will have to be resolved in future studies. In the light of the homology between MBPs and the chicken hepatic lectin discussed above, it is interesting that MBP is a hexamer, since the chicken protein has also recently been shown to be hexameric. The clustering of binding sites may be a characteristic feature of receptors which have high affinity for oligosaccharides with multiple terminal sugars.

Alternative Modes of Biosynthesis—It has recently been shown that the membrane-spanning segment of the rat asialglycoprotein receptor serves as an uncleaved, internal signal sequence which interacts with the signal recognition particle and directs this protein to be inserted into membranes in a transmembrane orientation with the COOH-terminal carbohydrate-recognition domain on the noncytoplasmic side of the membrane (44). In contrast, the MBPs are delivered to the noncytoplasmic side of the membrane by a more conventional cleaved, NH$_2$-terminal signal sequence. This indicates that the folding of the homologous carbohydrate-recognition domains in these two proteins is not likely to depend on a specific insertion pattern. This is consistent with recent evidence that the carbohydrate-recognition domain of the rat asialglycoprotein receptor is functional when it is directed to the lumen of microsomes by an NH$_2$-terminal, cleaved signal sequence derived from preproinsulin (43).

Comparison of MBPs with More Distantly Related Proteins

Proteins Containing Collagen-like Domains—The organization of the MBPs into collagenous and ligand-binding domains has parallels in other proteins which also have collagen-like regions. These include the asymmetric forms of acetylcholinesterase (35), complement protein Clq (34), and the apoprotein of pulmonary surfactant (45). The latter two examples are particularly informative. Each of these proteins consists of polypeptides which are linked by disulfide bonds that immediately precede collagenous domains at their NH$_2$ termini, whereas the COOH-terminal domain of each is non-collagenous.

In addition to this overall organizational homology, the dog pulmonary surfactant shows amino acid sequence homology with the MBPs (see Fig. 10). Pulmonary surfactant apoprotein and MBP-A can be aligned, with only three gaps, to show 30% identity in sequence. Seven gaps must be introduced to achieve an equivalent degree of identity with MBP-C. The strong homology between the MBPs and dog pulmonary surfactant apoprotein suggests that these proteins may be functionally related to each other. It may be particularly significant that virtually all of the "invariant" residues in the carbohydrate-recognition domains of the four carbohydrate-binding proteins shown in Fig. 10 are also found in the pulmonary surfactant apoprotein. This suggests that this protein may also have carbohydrate binding activity, a suggestion which can be tested experimentally. Alternatively, since the pulmonary surfactant apoprotein is known to have calcium ion-dependent lipid binding activity (46), it is possible that the binding sites for carbohydrates and phospholipids are structurally analogous. Finally, it is also interesting that the single interruption in the Gly-X-Y repeat pattern in all of these proteins falls in the same part of the collagenous domain, suggesting that this feature may have significance for the interaction of all of these proteins with a common effector protein.

The homology between the MBPs and complement Clq lies largely at the level of overall organization. Both of these proteins have short NH$_2$-terminal domains which are involved in interchain disulfide bond formation, followed by triple helix-forming domains. In each case, the pattern of Gly-X-Y repeats is interrupted near the middle. There are, however, distinct differences. The collagen-like portion of Clq is somewhat longer (25–27 Gly-X-Y repeats) than the corresponding domains of the MBPs. Also, the NH$_2$-terminal domains of the MBPs contain several cysteine residues, whereas the corresponding portions of Clq contain only 1. The COOH-terminal domain of each of these proteins is involved in a ligand binding activity—mannose-terminated oligosaccharides in the case of the MBPs and the Fc portion of immunoglobulins in the case of Clq. Although none of the conserved residues shown in Fig. 10 is present in Clq, and there is no clear homology between the two COOH-terminal domains, it is nonetheless intriguing that the fixation of complement is dependent on glycosylation of the Fc domain of the immunoglobulin (47).

The analogies between the MBPs and pulmonary surfactant apoprotein and complement of Clq may provide some indication of the natural function of the MBPs. It is possible that both the MBPs and pulmonary surfactant apoprotein participate in a primitive form of immune response. This would be analogous to the proposed function of the C-reactive protein, which recognizes the capsular polysaccharide of certain bacteria (5). Thus the presence of mannose might be taken as a relatively nontoxic marker for a bacterial surface in order to induce an effector function analogous to that induced more specifically by Clq. The presence of MBP in the circulation is consistent with this possibility. The pulmonary sur-
fiantant protein might serve a similar protective function at the surface of the lung.

**Relationship to Other Carbohydrate-binding Proteins in Serum**—C-reactive protein and serum amyloid proteins are two members of the pentraxin family, a group of pentameric and decameric proteins which show strong sequence homology with each other and which have characteristic binding activities (48). Comparison of the sequences of the MBPs with the pentraxins reveals that, as in the case of Clq, the conserved residues shown in the carbohydrate-recognition domains of Fig. 10 are not found in the pentraxins. The pentraxins contain a pair of highly conserved cysteine residues which are linked in a disulfide bond; there is no homology in their placement or surrounding residues with the 4 conserved cysteine residues in the carbohydrate-binding proteins shown in Fig. 10. Thus the carbohydrate-recognition activity exhibited by some members of the pentraxin family has a structural basis distinct from the family of proteins which contains the MBPs.

**Comparison with a Lectin from an Invertebrate**—A final comparison of the carbohydrate-recognition sequence in the MBPs sheds additional light on the evolution of this domain. A portion of the sequence of a galactose-binding protein from insect hemolymph (49) is shown aligned with the other sequences in Fig. 10. A clear homology with the carbohydrate-recognition activity exhibited of the invariant residues. This domain is found at the carbohydrate-recognition activity exhibited of the invariant residues. This domain is found at the

**Acknowledgments**—We thank Virginia McCreary, Judith Leung, and Edward Dunphy for assistance with DNA sequence analysis, and Dawn Farrell, John Hildreth, and Regina McStraw for aid in protein preparation and sequence analysis. We are also grateful to Jean Schwarzbauer, Richard Hynes, Michael McPhaul, and Thomas Sergat for providing cDNA and genomic libraries.

**REFERENCES**

1. Ashwell, G., and Harford, J. (1982) *Annu. Rev. Biochem.* 51, 531–554
2. Hoffmann, B., and Kornfeld, S. (1985) *J. Biol. Chem.* 260, 12098–12104
3. Birley, E. B., Gregory, W., Fletcher, P., and Kornfeld, S. (1979) *J. Cell Biol.* 81, 529–537
4. Barondes, S. H. (1984) *Science* 223, 1259–1264
5. Paeper, M. B., and Balskus, M. L. (1983) *Adv. Immunol.* 34, 141–211
6. Drickamer, K., Mamon, J. P., Binns, G., and Leung, J. O. (1984) *J. Biol. Chem.* 259, 770–778
7. Spiess, M., and Lodish, H. F. (1983) *Proc. Natl. Acad. Sci. U. S. A.* 82, 6465–6469
8. Chiacchio, K. B., and Drickamer, K. (1984) *J. Biol. Chem.* 259, 15440–15446
9. Kawasaki, T., Etoh, R., and Yamashina, I. (1978) *Biochem.* 6084
10. Mizuno, Y., Kozutsumi, Y., Kawasaki, T., and Yamashina, I. (1981) *J. Biol. Chem.* 256, 4247–4252
11. Townsend, R., and Staal, P. (1981) *Biochem.* 199, 209–214
12. Stahl, P. D., and Schleiszter, P. H. (1980) *Trends Biochem. Sci.* 21, 194–196
13. Maynard, Y., and Benzing, J. U. (1982) *J. Biol. Chem.* 257, 3783–3794
14. Mori, K., Kawasaki, T., and Yamashina, I. (1983) *Arch. Biochem. Biophys.* 222, 542–552
15. Kozutsumi, Y., Kawasaki, T., and Yamashina, I. (1981) *J. Biochem.* (Tokyo) 90, 1759–1807
16. Kawasaki, N., Kawasaki, T., and Yamashina, I. (1983) *J. Biochem.* (Tokyo) 94, 937–947
17. Forstdet, N., and Porath, J. (1975) *FEBS Lett.* 57, 187–191
18. Forstdet, N. (1974) *Methods Enzymol.* 34, 13–30
19. Drickamer, K. (1981) *J. Biol. Chem.* 256, 5827–5839
20. Schwarzbaumer, J., Tannen, J. W., Lemischka, I. R., and Hynes, R. O. (1983) *Cell* 33, 421–431
21. Okuyama, H., and Berg, P. (1983) *Mol. Cell. Biol.* 3, 280–289
22. Heinrikson, R. L., and Meredith, S. C. (1984) *Anal. Biochem.* 136, 65–74
23. Laemmli, U. K. (1970) *Nature* 227, 680–685
24. Tarr, G. E. (1975) *Anal. Biochem.* 63, 361–370
25. Smithies, O., Gibson, D., Fanning, E. M., Goodfellow, R. M., Gilman, J. G., and Ballantyne, D. L. (1971) *Biochemistry* 10, 4912–4921
26. Rose, S. M., and Schwartz, B. D. (1980) *Anal. Biochem.* 107, 206–213
27. Rittenhouse, J., Chatterjee, T., Marcus, F., Reardon, L., and Heinrikson, R. L. (1983) *J. Biol. Chem.* 258, 7648–7652
28. Holland, E. C., Leung, J. O., and Drickamer, K. (1984) *Proc. Natl. Acad. Sci. U. S. A.* 81, 7388–7392
29. Grunstein, M., and Hoggess, D. S. (1975) *Proc. Natl. Acad. Sci. U. S. A.* 72, 3961–3965
30. Mannitis, T., Fritsch, E. F., and Sambrook, J. (eds) (1982) *Molecular Cloning: A Laboratory Manual*, pp. 309–328, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
31. Leung, J. O., Holland, E. C., and Drickamer, K. (1985) *J. Biol. Chem.* 260, 12523–12527
32. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503–517
33. Bornstein, P., and Traub, W. (1975) in *The Proteins* (Nureith, H., and Hill, E. L., eds), 3rd ed., vol. 4, pp. 611–632, Academic Press, New York
34. Reid, K. B. M. (1983) *Biochem. Soc. Trans.* 11, 1–12
35. Rosenberry, T. L., and Richardson, J. M. (1977) *Biochemistry* 16, 3590–3598
36. Sengist, R., and Harper, E. F. (1971) in *The Enzymes* (Boyce, P., D., ed) 3rd ed., Vol. 3, pp. 649–697, Academic Press, New York
37. Lathe, R. (1985) *J. Mol. Biol.* 183, 1–12
38. Spiero, R. G. (1973) *Adv. Protein Chem.* 27, 349–467
39. von Heijne, G. (1973) *Biochemistry* 24, 184–190
40. von Heijne, G. (1974) *Biochemistry* 13, 245–251
41. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Biochem.* 50, 349–383
42. von Heijne, G. (1983) *Eur. J. Biochem.* 133, 17–21
43. Haue, E. C., Holland, E. C., Carrera, G. M., and Drickamer, K. (1986) *J. Biol. Chem.* 261, 4940–4947
44. Holland, E. C., and Drickamer, K. (1985) *J. Biol. Chem.* 251, 1286–1292
45. Benson, B., Hawgood, S., Schilling, J., Clements, J., Damm, D., Cordell, B., and White, R. T. (1985) *Proc. Natl. Acad. Sci. U. S. A.* 82, 6579–6583
46. Hawgood, S., Benson, B. J., and Hamilton, R. L., Jr. (1985) *Biochemistry* 24, 184–190
47. Nose, M., and Wiegell, H. (1983) *Proc. Natl. Acad. Sci. U. S. A.* 80, 6632–6636
48. Pelli, F., Prus, M., and Frangione, B. (1985) *J. Biol. Chem.* 260, 12985–12988
49. Takashashi, K., Komano, H., Kawaguchi, N., Kitamura, N., Nakashita, S., and Natori, S. (1985) *J. Biol. Chem.* 260, 12228–12235
50. Komano, H., Nozawa, R., Mizuno, D., and Natori, S. (1983) *J. Biol. Chem.* 258, 2143–2147
Structure of Mannos-binding Proteins

EXPERIMENTAL PROCEDURES

Materials — Collagenase (prepared from Clostridium histolyticum; Code CLS25A) was a product of Worthington Biochemicals. 1,2-Diiodoacetate was from American Biochemicals. V8 protease, alkaline phosphatase, and nitrocellulose were obtained from Sigma. Mannose-Seraphin was prepared by coupling of mannose to Sepharose 4B with dicyclohexylcarbodiimide (DCC) in dimethylformamide. Other materials for protein chemistry were obtained from the sources previously indicated (16,19). Restriction enzymes were from New England Biotools. TSK DNA ligase and polynucleotide kinase were obtained from Boehringer Mannheim. Two rat liver enzymes (20) were kindly provided by Jean Schwarzbauer and Richard Myers (Massachusetts Institute of Technology), while a library derived from rat pancreas was a generous gift of Robert McPhail (Dartmouth Medical School). A rat genomic library provided by Thomas Sargent (National Institutes of Health) was also utilized.

Preparation of MBP — Two procedures were employed for the preparation of MBP. The first (21) was that of a native enzyme. In the first method, based in that of Kawasaki et al (16,19), MBP was prepared from a solution of invertase-Seraphin from a strain of E. coli 055:B5. Invertase, Seraphin, and a mouse anti-seraphin antibody were detected in a gel column of galactoside-Sepharose. The crude MBP was purified by passage through a second gel column of phosphoethanolamine-Sepharose. The purified MBP was then isolated by centrifugation on an agarose column from a solution of invertase-Sepharose columns. Fractions containing MBP were assayed by binding of 125I labeled (1) and by SDS-polyacrylamide gel electrophoresis. MBP was also prepared from an inverted column of immobilized phosphoethanolamine-Sepharose, by the procedure described previously (16). In certain cases, precipitation was first performed with 8.5 M urea, 0.5 M NaCl, and 100 mM Tris (pH 8.3) to 0.1 M NaCl, 10 mM Tris (pH 8.3). The material remaining in the supernatant was recovered by lyophilization. Other procedures were identical to those described previously (16,19). Rat liver enzymes were also used. Rat liver enzymes were also used.

Degradation and isolation of (Mannose)-phosphoethanolamine-Sepharose columns were prepared with 200 ng of 5 M dithiothreitol for 3 min at 37°C. The resulting solution of proteolysis was then excluded against distilled water. The resulting solution was added to a gel column of 5 M Tris (pH 8.3), 0.5 M NaCl, and 100 mM Tris (pH 8.3). The material remaining in the supernatant was recovered by lyophilization. Other procedures were identical to those described previously (16). In certain cases, precipitation was first performed with 8.5 M urea, 0.5 M NaCl, and 100 mM Tris (pH 8.3) to 0.1 M NaCl, 10 mM Tris (pH 8.3). The material remaining in the supernatant was recovered by lyophilization. Other procedures were identical to those described previously (16,19). Rat liver enzymes were also used.

Isoelectric focusing — A sample buffer containing 5 M urea and 10 M NaCl was added to the sample buffer at a concentration of 1 M (v/v). The sample buffer was obtained by centrifuging on an agarose column from a solution of invertase-Sepharose columns. Fractions containing MBP were assayed by binding of 125I labeled (1) and by SDS-polyacrylamide gel electrophoresis. MBP was also prepared from an inverted column of immobilized phosphoethanolamine-Sepharose, by the procedure described previously (16). In certain cases, precipitation was first performed with 8.5 M urea, 0.5 M NaCl, and 100 mM Tris (pH 8.3) to 0.1 M NaCl, 10 mM Tris (pH 8.3). The material remaining in the supernatant was recovered by lyophilization. Other procedures were identical to those described previously (16,19). Rat liver enzymes were also used.

N-terminal amino acid analysis — N-terminal amino acid analysis was performed by Edman degradation, using an Applied Biosystems 890A analyzer. The elution of each amino acid was monitored using a Murakami Chromatography Analysis System. The elution of each amino acid was monitored using a Murakami Chromatography Analysis System. The elution of each amino acid was monitored using a Murakami Chromatography Analysis System.

S-hydroxylation of peptides — S-hydroxylation of peptides was performed by incubation of reduced and alkylated MBP with 1.0% H2O2 in 0.05% EDTA (pH 7.8) for 2 h at 37°C. The reaction mixture was then centrifuged, and the supernatant was monitored using an Applied Biosystems 890A analyzer. The elution of each amino acid was monitored using a Murakami Chromatography Analysis System. The elution of each amino acid was monitored using a Murakami Chromatography Analysis System.

Nuclear acid chemistry — Preparation and characterization of oligonucleotide probes, preparation of plasmid DNA, 11] labeling of DNA sequence fragments, and sequencing of DNA by the chemical degradation method were performed as previously described (18).

Isolation of DNA and genomic clones — Screening of the plaque DNA library was performed at a density of 500,000 plaques per 25 cm plate (9) following the method of Crowther et al (21). The library was screened by using 1.0% agarose gel electrophoresis. Duplicate nitrocellulose filter lifts were prehybridized in a solution containing 5% dextran sulfate, 0.5% SDS, 70% formamide, 0.1 M sodium phosphate, pH 6.8, 0.5 M EDTA, 5 × Denhardt's solution (NS), and 0.1% SDS for 2 h at 45°C. Hybridization probe (200 ng/ml) was added, and the hybridization mixture was prehybridized for 2 h at 45°C. The filters were washed twice for 25 min at 55°C in a solution of 0.5% SDS, 0.1 M sodium phosphate, pH 6.8, and 0.5 M EDTA. Following plaque purification of positive plaque, the inserts were excised by digestion with EcoRI and added into the blue side of phage bluse (19). Detergent was added to the solution containing 3% SDS, 5% sodium dodecyl sulfate, and 1 M NaCl. Digestion with EcoRI and added into the blue side of phage bluse (19). Detergent was added to the solution containing 3% SDS, 5% sodium dodecyl sulfate, and 1 M NaCl.

Schematic diagram — The schematic diagram of the DNA library utilized in this study is shown in Fig. 5. The DNA library was constructed by ligating the DNA fragments into pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18). Relevant portions of the genomic clone of E. coli MBP were identified by hybridization analysis of restriction digests of plasmid DNA (21). A HindIII fragment which hybridized with the 3' end of the DNA clone was subcloned into the HindIII site of pBR322 and screened as previously described (18).

RESULTS

Fig. 2. (Left) Separation of tryptic peptides of MBP. A preparation of MBP was digested enzymatically with Kex2 and the resulting tryptic material was reduced and carboxymethylated. This fraction was enriched in MBP-A relative to the original preparation. Approximately 2 mg of this material was subjected to tryptic digestion. The resulting peptides were separated by chromatography on a Waters 1-L25 column eluted with 0.1% acetic acid (pH 3.5). The portion of the eluate indicated by the horizontal bar was pooled, dried, and separated by electrophoresis and chromatography on thin layer cellulose plates (5.83). 10 μg of each fraction was visualized by staining with fluorescamine and subjected to Edman degradation in the sequencing gel (see Table III).

Fig. 4. (Right) Isolation of a large chloroplast fragment of MBP. A preparation of MBP was subjected to collagenase digestion, and the resulting collagenase-resistant fragment was isolated by affinity chromatography. This fraction (approximately 0.5 mg) was further digested with trypsin. The resulting peptides were separated by chromatography on a Waters 1-L25 column eluted with 0.1% acetic acid (pH 3.5). The portion of the eluate indicated by the horizontal bar was pooled and further digested by Edman degradation on a Beckman Ultrorac 42-CD column eluted with a linear gradient of acetonitrile in the presence of 0.1% trifluoroacetic acid (pH 3.5). The portion of the eluate indicated by the horizontal bar was pooled and further digested by Edman degradation on a Beckman Ultrorac 42-CD column eluted with a linear gradient of acetonitrile in the presence of 0.1% trifluoroacetic acid (pH 3.5). The portion of the eluate indicated by the horizontal bar was pooled and further digested by Edman degradation on a Beckman Ultrorac 42-CD column eluted with a linear gradient of acetonitrile in the presence of 0.1% trifluoroacetic acid (pH 3.5).
Fig. 6. Sequencing strategies for clones pMBP-A and pMBP-C. Restriction restriction sites in the inserts of each plasmid are marked at the top. Sequences shown were labeled at their 5' termini. Portions of fragments for which sequences were read are denoted by white boxes. In both cases, greater than 95% of each sequence was determined for both strands. The partial clone of MBP-C which was initially isolated covers the region indicated as Pm; this restriction fragment was used as a probe to isolate additional MBP-C clones and the MBP-A clones. In the case of pMBP-A, the fragment used as a probe to identify clones of the gene encoding this protein (Fig. 5) is also marked.

Fig. 7. Sequencing of pMBP-C insert and peptides derived from MBP-C. The nucleotide sequence of the cloned insert is translated to yield the amino acid sequence of MBP-C. Portions of the insert are compared to the sequence described at the N-terminal end of the protein and the sequence of tryptic peptides derived from this form of MBP. Asterisks denote stop codons.

TABLE I

| Amino acid sequence of predicted and experimentally determined MBP | Yield of each amino acid residue estimated by high pressure liquid chromatography (in pmol) | Precision MBP-P yielded one sequence, while Fraction MBP-P yielded two sequences. |
|---|---|---|
| Glu| 156| 125| Ser| 64|
| Thr| 72| 73| Tyr| 129|
| Leu| 280| 243| Ser| 76|
| Ser| 74| 103| Glu| 116|
| Gly| 162| 296| Thr| 39|
| Ala| 122| 134| Glu| 66|
| Val| 167| 148| Cys| 21|
| Ser| 37| 13| Leu| 74|
| Pro| 23| 25| Lys| 44|
| Val| 12| 11| Thr| 20|
| Pro| 9| 13| Cys| 19|
| Leu| 15| 17| Ser| 9|
| Ala| 16| 18| Cys| 13|
| Ser| 17| 18| Ser| 22|

Cytochrome residues were confirmed by determination of radioactivity in each cycle.
TABLE II

Amino acid sequences of tryptic peptides derived from MBP

| CYCLE | PEPTIDE 1A | PEPTIDE 1B | PEPTIDE 3 | PEPTIDE 4 | PEPTIDE 5 | PEPTIDE 6 | PEPTIDE 7 | PEPTIDE 8 | PEPTIDE 9A | PEPTIDE 9B |
|-------|------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|       | 1 nmol     | 1 nmol     | 5 nmol    | 0 nmol    | 3 nmol    | 13 nmol   | 6 nmol    | 8 nmol    | 6 nmol    |
| 1     | Lys 0.6    | Ser 0.2    | Asn 2.9   | Lys 1.9   | Gly 4.5   | Val 3.5   | Ala 6.2   | Ser 0.2   |           |
| 2     | Ser 0.2    | Gly 0.6    | Ala 3.8   | Lys 1.6   | Leu 4.8   | Arg 1.3   | Ile 6.0   | Lys 1.5   |           |
| 3     | Gly 0.3    | Lys 0.5    | Gly 2.3   | Gly 1.8   | Glu 4.8   | Glu 2.2   | Leu 0.8   |           |           |
| 4     | Lys 0.2    | Gly 1.2    | Glu 1.0   | Gly 3.1   | Asn 3.2   | Glu 1.1   |           |           |           |
| 5     | Asn 1.8    | Lys 1.3    | Pro 3.1   | Val 2.9   | Leu 1.5   |           |           |           |           |
| 6     | Arg 0.3    | Lys 1.7    | His 2.0   | Ala 3.6   | Thr 0.3   |           |           |           |           |
| 7     | Gly 2.4    | Lys 2.0    | Asn 1.0   |           |           |           |           |           |           |
| 8     | Lys 3.5    | Lys 1.8    |           |           |           |           |           |           |           |

TABLE III

Amino acid sequences of tryptic, collagenase, and chymotryptic peptides derived from MBP

| CYCLE | PEPTIDE T-2 | PEPTIDE C-1 | PEPTIDE Cl-1 |
|-------|-------------|-------------|--------------|
|       | 5 nmol      | 1 nmol      | 1.5 nmol     |
| 1     | Leu 1.9     | Gly 0.8     | Ala 0.8      |
| 2     | Gly 1.2     | Pro 0.36    | Ile 0.67     |
| 3     | Pro 0.8     | Lys 0.07    | Glu 0.78     |
| 4     | Hpr 0.9     | Gly 0.23    | Asn 0.87     |
| 5     | Gly 2.1     | Asp 0.19    | Val 0.63     |
| 6     | Ser 0.2     | Arg 0.19    | Ala 0.67     |
| 7     | Val 1.2     | Gly 0.20    | Lys 0.82     |
| 8     | Gly 1.0     | Gly 0.20    | Asp 0.26     |
| 9     | Ala 1.2     | Ser 0.07    | Val 0.39     |
| 10    | Hpr 0.4     | Val 0.13    | Ala 0.42     |
| 11    | Gly 1.0     | Glu 0.11    | Phe 0.41     |
| 12    | Ser 0.1     | Phe 0.15    | Leu 0.69     |
| 13    | Glu 0.5     | Asp 0.09    | Gly 0.39     |
| 14    | Gly 0.8     | Thr 0.06    | Ser 0.38     |
| 15    | Pro 0.1     | Thr 0.07    | Thr 0.12     |
| 16    | Arg 0.1     | Asp 0.10    | Arg 0.18     |
| 17    | Gly 0.4     | Le 0.09     | Glu 0.30     |
| 18    | Glu 0.2     | Asp 0.04    | Arg 0.31     |
| 19    | Leu 0.09    | Thr 0.07    |             |
| 20    | Gly 0.4     | Glu 0.15    |             |
| 21    | Arg 0.1     | Le 0.06     | Asn 0.20     |
| 22    | Arg 0.1     | Ala 0.08    | Val 0.19     |
| 23    | Ala 0.14    | Phe 0.32    |             |
| 24    | Leu 0.12    | Glu 0.09    |             |
| 25    | Ser 0.04    | Leu 0.10    |             |
| 26    | Thr 0.06    |             |             |
| 27    | Leu 0.07    | Glu 0.11    |             |
| 28    | Ala 0.05    |             |             |