Protein Science (2000), *9*:1607–1617. Cambridge University Press. Printed in the USA.

REVIEW

Collectin structure: A review

KJELL HÅKANSSON[1] AND KENNETH B.M. REID[2]

[1]Department of Microbiology, University of Illinois at Urbana-Champaign, B103 CLSL, 601 S. Goodwin Avenue, Urbana, Illinois 61801
[2]MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom

(RECEIVED May 17, 2000; FINAL REVISION June 21, 2000; ACCEPTED June 22, 2000)

Abstract

Collectins are animal calcium dependent lectins that target the carbohydrate structures on invading pathogens, resulting in the agglutination and enhanced clearance of the microorganism. These proteins form trimers that may assemble into larger oligomers. Each polypeptide chain consists of four regions: a relatively short N-terminal region, a collagen like region, an α-helical coiled-coil, and the lectin domain. Only primary structure data are available for the N-terminal region, while the most important features of the collagen-like region can be derived from its homology with collagen. The structures of the α-helical coiled-coil and the lectin domain are known from crystallographic studies of mannan binding protein (MBP) and lung surfactant protein D (SP-D). Carbohydrate binding has been structurally characterized in several complexes between MBP and carbohydrate; all indicate that the major interaction between carbohydrate and collectin is the binding of two adjacent carbohydrate hydroxyl group to a collectin calcium ion. In addition, these hydroxyl groups hydrogen bond to some of the calcium amino acid ligands. While each collectin trimer contains three such carbohydrate binding sites, deviation from the overall threefold symmetry has been demonstrated for SP-D, which may influence its binding properties. The protein surface between the three binding sites is positively charged in both MBP and SP-D.

Keywords: CRD; C-type lectin; MBP; SP-A; SP-D; three-dimensional structure

The collectins are one of the groups of calcium dependent or C-type lectins found in mammals. They represent a nonclonal and innate host defense system, which is functional in the absence of, and prior to, the development of the adaptive antibody based immune system. Six different collectins have been characterized as extracellular proteins from three major locations. Lung surfactant proteins A and D (SP-A, SP-D) are found in the pulmonary surfactant on the epithelial lining of the lungs where they provide protection against invading pathogens and allergens. Mannan-binding proteins (MBP-A, MBP-C) are found both in serum and in the liver. The serum protein (MBP-A) differs from the hepatic form in some species, but not in man. Conglutinin (BC) and collectin 43 (CL-43) are bovine serum proteins more closely related to SP-D than to MBP. In addition, an intracellular collectin of unknown function, collectin liver 1 (CL-L1), was recently discovered in liver cells (Ohtani et al., 1999).

Reprint requests to: Kjell Håkansson, Department of Microbiology, University of Illinois at Urbana-Champaign, B103 CLSL, 601 South Goodwin Avenue, Urbana, Illinois 61801; e-mail: kjell@scs.uiuc.edu.

Abbreviations: CRD, carbohydrate recognition domain; SP-A, SP-D, lung surfactant protein A and D; MBP, MBP-A, MBP-C, mannan-binding protein A and C; CL-43, collectin 43; CL-L1, collectin liver 1.

The collectins target invading pathogens—mainly bacteria but also fungi, viruses, and potential allergens—by binding to their surface carbohydrates. This encounter usually results in aggregation or agglutination of the target particle, and any further subsequent biological response depends upon the location and collectin involved. Most collectins interact with and stimulate phagocytic cells, followed by opsonization of the target. In addition, serum mannan-binding protein is able to trigger the classical complement system, via two newly identified serine proteases (MASP-1 and MASP-2) (Matsushita & Fujita, 1992; Thiel et al., 1997), ultimately resulting in lysis of the invading microorganism.

The name collectin is derived from the words "collagen" and "lectin." These molecules are composed of four different regions: an N-terminal region, a collagenous region, an α-helical coiled-coil, and a C-terminal lectin domain. With the exception of the N-terminal region, these regions or domain types are also found in molecules other than collectins (Fig. 1). The presence of a collagen-like region in these molecules imposes a trimeric structure. This is a prerequisite for their proper function; the carbohydrate affinity of a single collectin carbohydrate recognition domain (CRD) is weak but the trimeric organization permits a trivalent and hence stronger interaction between collectin and carbohydrate-containing target surface. In most of the collectins, these trimers are further assem-

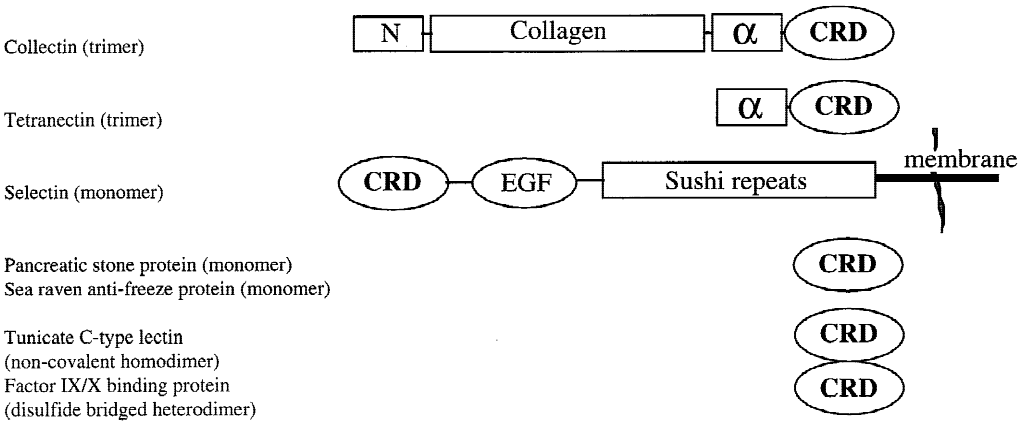

Fig. 1. Schematic representation of some proteins containing the C-type lectin fold. The structures are known, to atomic resolution, for all the listed proteins. CRD, carbohydrate recognition domain; EGF, epidermal growth factor domain.

bled into larger entities, which enables them to cross-link several target particles and perhaps also to interact simultaneously with target and with host cells.

Lung surfactant protein D (SP-D) and bovine conglutinin are X-shaped molecules consisting of four trimers extending pairwise from the hub with the lectin domains at the tip of the collagenous arms (Crouch et al., 1994). These dodecamers have been shown to associate further into even larger complexes, with an increased capacity to aggregate bacteria. Lung surfactant protein A (SP-A) (Voss et al., 1988) and most mannan binding proteins (Lu et al., 1990) form bouquet-like complexes with up to six trimers. The degree of oligomerization differs between the different proteins, but may also vary within the molecular population of a single preparation. Only the larger complexes, i.e., pentamers or hexamers, give a full biological response in terms of macrophage stimulation and complement activation. Interestingly, the overall structural organization of SP-A and the oligomeric MBPs is reminiscent of that of C1q of the complement system, extending the analogy between these proteins to encompass both function and structure. However, the degree of oligomerization of MBP seems to be lower in the liver than in serum, and rat MBP-C has been reported to be a single trimer (Wallis & Drickamer, 1997) like collectin 43 (CL-43) (Holmskov et al., 1995).

There are a number of excellent reviews describing the biological, physiological, and functional properties of the collectins (Hoppe & Reid, 1994; Weis & Drickamer, 1996; Lu, 1997; Crouch, 1998). The present review will focus on the structural aspects of the mannan-binding proteins and the lung surfactant proteins.

N-terminal region

The N-terminal region is defined as the segment N-terminal to the first collagenous triple-helix residue. This relatively cysteine-rich region stabilizes the trimers through disulfide bridging (Crouch et al., 1994; Holmskov et al., 1995) and links them together in the collectin oligomers. There seems to be no overall homology between the collectins in this part of the molecule. The only general trend (to which canine SP-A, Rhesus monkey MBP-C, and CL-L1 do not conform) seems to be two cysteine residues separated in sequence by 4–5 amino acids, and perhaps also that the amino acid preceding the second of these cysteines is hydrophobic in most of the collectins. Most of the other amino acids separating these two cysteine residues are relatively hydrophobic in MBP and SP-D but hydrophilic in SP-A.

The collectins can nonetheless be divided into four different groups with related N-terminal regions. The N-terminal sequences of SP-D, conglutinin, and collectin-43 are clearly related and of similar length, i.e., 25–28 amino acids (Fig. 2), and include two cysteine residues. The N-terminal region of SP-A is much shorter and, due to variation in signal peptidase cleavage, is not homogenous; there are two isoforms with seven or ten residues containing one or two cysteines, respectively (Elhalwagi et al., 1997). The N-terminal asparagine (position 4 in the longer isoform) of rat SP-A is glycosylated (McCormack et al., 1994). The third group consists of mannan-binding proteins A and C, which have distinct, yet related, N-terminal regions, each approximately 20 amino acids long. The hepatic forms rat MBP-C and mouse MBP-C have only two cysteines while the other mannan-binding proteins have three, although the sequence of rhesus monkey MBP-C deviates from the similarity displayed by the other sequences. Human MBP has three cysteines despite its closer evolutionary relationship to rat MBP-C and forms larger oligomers in serum than it does in the liver (Kurata et al., 1994). Comparisons of hybrid peptide chains of rat MBP-A and MBP-C and site-directed mutagenesis led to the conclusion that oligomerization depends on the N-terminal region and that the most N-terminal of its cysteines (Cys6) is indispensable for the formation of oligomers of trimers. The other two cysteines (Cys13 and Cys18) are involved in intratrimer disulfide bridge formation (Wallis & Drickamer, 1999). These two cysteines and the corresponding residues in rat MBP-C and collectin 43 form an asymmetric A1–B1 B2–C2 C1–A2 disulfide pattern (Rothmann et al., 1997; Wallis & Drickamer, 1997, 1999). Formation of SP-A oligomers is also dependent on the N-terminal region, but neither of its two cysteines are individually required for oligomerization (McCormack et al., 1997, 1999; Zhang et al., 1998). Accordingly, one of these cysteines is naturally replaced by an asparagine in canine SP-A. However, it is reasonable to assume that, as with the other collectins, disulfide bridging is an additional stabilizing factor. There are no free SH groups, and reduction of SP-D destabilizes the trimer as well as the tetramer of trimers (Crouch et al., 1994). The fourth group, so far represented only by human CL-L1, has a much longer N-terminal sequence with only one cysteine.

```
                                1
Rat SP-D                        AEMKTLSQRS ITNTCTLVLC SPTENGLP
Mouse SP-D                      AEMKSLSQRS VPNTCTLVMC SPTENGLP
Conglutinin                     AEMTTFSQKI LANACTLVMC SPLESGLP
Bovine SP-D                     AEMKIYSQKT MANACTLVMC SPPEDGLP
Human SP-D                      AEMKTYSHRT MPSACTLVMC SSVESGLP
Porcine SP-D                    AEMKTYSQRA VANACALVMC SPMENGLP
Bovine collectin-43             EEMDVYSEKT LTDPCTLVVC APPADSLR
Consensus                        EM   S        C LV C       L

                                1
Mouse SP-A                   IKC NGTEVCAGSP
Rat SP-A                     IKC NVTDVCAGSP
Rabbit SP-A                 SDTC DTKDVCIGSP
Guinea Pig SP-A              IMC NRTDFCVGSP
Canine SP-A                  IEN NTKDVCVGNP
Human SP-A                   AAC EVKDVCVGSP
Porcine SP-A                 LEC DVKEVCLASP
Consensus                          C   P


Mouse MBP-A                   SS GSQTCEDTL KTC.SVIAC
Rat MBP-A                      S GSQTCEETL KTC.SVIAC
Rhesus MBP-A                     ETKACEDAQ KTC.SVITCG IPVTN
Mouse MBP-C                      ETLT.EGVQ NSCP.VVTCS SPGLN
Rat MBP-C                        ETLT.EGAQ SSCP.VIACS SPGLN
Rhesus MBP-C          CLSFICTN STGGPQPHA DGSPEVIACN SPGIN
Human MBP                        ETVTCEDAQ KTCPEVIACN SPGIN
Porcine MBP-C                    ETENCEDIQ NTCL.VISCD SPGIN
Bovine MBP-C                     ETENCENIR KTCP.VIACG PPGIN
Consensus                                      V  C


Human CL-L1          MNGFASLLRR NQFILLVLFL LQIQSLGLDI DSRPTAEVCA THTISPGPK
```

Fig. 2. Sequences of the N-terminal region of some collectins. Most mature SP-A sequences are reported to start at residue 1, but variants starting from residue −3 are known.

Collagenous region

Despite their abundance in the animal organism, there are no experimentally determined atomic resolution structures of native collagens. Our knowledge of collagen structure is mainly derived from fiber diffraction studies, model building (for a review see Beck & Brodsky, 1998), and the crystallographic structures of short model compounds that have been solved recently (Bella et al., 1994; Kramer et al., 1998, 1999). A collagen structure can be recognized from the amino acid sequence with its characteristic Gly-X-Y repetitive pattern, where X and Y can be any amino acid but are frequently prolines or hydroxyprolines. Each of the three chains forms a left-handed polyproline II like helix and are coiled around each other in a right-handed manner with the glycine residues in the interior of the superhelix. Interchain hydrogen bonds between N-H groups of glycine and the C=O groups of the amino acid in X position stabilize the structure, and since there are no intrachain hydrogen bonds the collagen helix can exist only as a trimer. The helical parameters and conformational details, such as helical twisting, are sequence dependent and differ between imino acid rich and amino acid rich regions (Kramer et al., 1999). The collagen triple helix is surrounded by a network of water molecules that interacts with most polar groups and all residues are exposed to the solvent, either by their side chains or carbonyl groups. This hydration is further enhanced by hydroxylation of some of the prolines in the Y position without which collagen is unstable at biological temperature (Kramer et al., 1998). Accordingly, the triple helix in the collectins is rich in charged amino acids. It has been pointed out (Wallis & Drickamer, 1997) that the staggered geometry of the triple helix introduces asymmetry in the N-terminal region of the collectins in agreement with the asymmetric disulfide bond pattern.

The triple helix appears to be ideal as a cross-linker between different domains due to its high tensile strength, stability, and relative resistance to proteolysis. In addition, a triple helical region can also mediate binding interactions with other macromolecules as has been shown for C1q and the macrophage scavenger receptor. The collagenous region in MBP contains the binding site for its associated serine proteases (MASP-1 and MASP-2) through which the complement cascade of reactions is triggered (Matsushita &

Table 1. *Some characteristics of the collectin N-terminal and the collagenous regions*

	N-terminal region		Collagenous region		
	No. of residues	No. of cysteines	No. of GXX triplets	No. of imino acids[a]	No. of helix interruptions
Human SP-A	7–10	1–2	23	20	1 PCPP insertion
Rat MBP-A	18	3	17	8	1 GQ insertion
Human MBP	21	3	19	12	1 GQ insertion
Rat MBP-C	19	2	19	10	1 GQ insertion
Human SP-D	25	2	59	30	0
Conglutinin	25	2	56	32	1 CPH insertion
Collectin 43	28	2	38	24	0
Collectin liver 1	46	1	24	6	0

[a]Includes prolines within interruptions and immediately preceding the first GXX triplet.

Fujita, 1992; Thiel et al., 1997). SP-A and presumably also the other collectins bind to their macrophage receptor through the collagenous region (Malhotra et al., 1993). As can be seen in Table 1, the length of the collagenous region differs between the collagens; the X-shaped SP-D and bovine collectin have the longest triple-helical regions, and SP-A and MBP the shortest. The percentage of imino acid in the collagenous region varies between 16 and 29% for the extracellular collagens, which is similar to what is found in collagen, whereas the imino acid content in CL-L1 is only 8%. Most of the collectin prolines in the Y position are hydroxylated. Hydroxylated and glycosylated lysines have been demonstrated in SP-D, MBP, collectin 43 and conglutinin (Colley & Baenziger, 1987; Sano et al., 1987; Persson et al., 1989; Lee et al., 1991; Rothmann et al., 1997), whereas the post-translational fate of the SP-A and CL-L1 lysines remains to be investigated. Some proteins contain an interruption of the Gly-X-Y sequence (Table 1). In the case of SP-A, there appears to be a kink in the collagenous rod in this part of the helix, as judged by electron microscopy. The collagen triple helices aggregate laterally at their N-terminal ends, but there is no evidence for covalent cross-linking. The nature of this intertrimeric interaction is not known but is likely to involve a network of hydrogen bonded water molecules, as has been observed in the crystal structure of collagenous peptides (Bella et al., 1994; Kramer et al., 1998). At the site of the Gly-X-Y interruption, the helices make a 60° bend and diverge (Voss et al., 1988) in a manner familiar from studies on the C1q structure (Lu et al., 1993). Oligomerization of MBP-A is in part promoted by the N-terminal part of its collagenous region (Wallis & Drickamer, 1999), but this region is not indispensable for SP-A oligomer formation (McCormack et al., 1997).

α-Helical coiled-coil region

Right handed α-helices twisting around each other in a left-handed manner are more ubiquitous in their occurrence than the animal collagen triple helix. Unlike the staggered triple helix, the side chains from the different polypeptide chains are in register and the junction between the collagenous region and the coiled-coil, therefore, cannot be symmetric. The coiled-coil has no absolute amino acid sequence requirement, although proline residues, unable to form amino hydrogen bonds, usually introduce bends in α-helices. However, the coiled-coil is stabilized by hydrophobic amino acids at every turn of the helix. These hydrophobic amino acids form the interior of the coiled-coil. An α-helical coiled-coil can often be recognized from the amino acid sequence by its characteristic heptad repeat pattern a-b-c-d-e-f-g, where residues "a" and "d" are hydrophobic amino acids. The structure of several coiled-coils, including the structure of trimeric MBP and SP-D fragments comprising the coiled-coil region and the lectin domains (Sheriff et al., 1994; Weis & Drickamer, 1994; Håkansson et al., 1999) and the related tetranectin molecule (Nielsen et al., 1997), are known from X-ray crystallography. The residues have been aligned according to their positions in the three-dimensional (3D) structure in Figure 3. The α-helices make 8–9 turns and there are several violations against the heptad rule, i.e., hydrophilic and even charged residues can be found in some of the "a" and "d" positions. The heptad pattern is further obscured by position "g" in the fourth and position "c" in the fifth heptad, which are involved in hydrophobic interactions with the lectin domain of the collectins. In tetranectin, one of these positions is occupied by a cysteine residue that is disulfide bonded to the C-type lectin domain. The helical region in SP-D is abruptly terminated by a proline residue with a cis peptide bond.

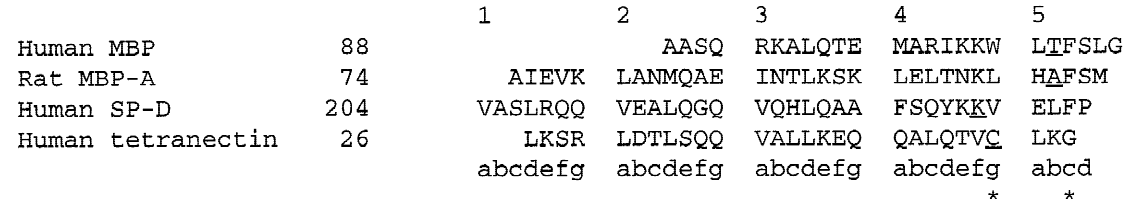

Fig. 3. Alignment of the coiled-coil regions of human and rat mannan binding proteins, human lung surfactant protein D and tetranectin based on the 3D structures. Most of the residues in positions "a" and "d" of the heptad repeat are hydrophobic although there are several violations against this rule. In human MBP, it is Leu110 ("5e") in the fifth heptad repeat rather than Ser109 ("5d") that makes the helix–helix contact. The main-chain Cα of residues "1a" in hSP-D and "2d" in hMBP and the side chain of residue "3a" in hMBP are actually exposed but this is probably due to their proximity to N-termini of the fragments used. The last amino acid, *i*, in each helix that makes a hydrogen bond with residue $i + 4$ is underlined. Residues "4g" and "5c" (marked with an asterisk (*)) belong to a hydrophobic cluster that are in contact with the lectin domain. In tetranectin, there is a disulfide bridge between Cys50 ("4g") and Cys60 of the lectin domain.

C-type lectin domain

Many evolutionarily related sequences belonging to the C-type (calcium dependent) lectin family are known, but they show no overall sequence similarity. However, they can usually be classified on the basis of four cysteine residues, being involved in disulfide bridging, which are the trademark of this domain type. The framework surrounding this domain type can be very different from its C-terminal location in the collectins. The monomeric and membrane-bound selectins have an N-terminal C-type lectin domain that mediates adhesion between certain cell types through carbohydrate binding. Other proteins, such as phospholipase A2 receptor and the macrophage mannose receptor, contain multiple copies of C-type lectin domains within a single polypeptide. Proteins possessing this domain type are not necessarily lectins. Type II antifreeze protein present in some arctic fish and the mammalian pancreatic stone protein bind to and inhibit the growth of ice and calcium carbonate crystals, respectively. Some examples of proteins with a C-type lectin domain are schematically represented in Figure 1.

The 3D structure of the lectin domain is known to atomic resolution from X-ray analyses of rat and human mannan-binding protein fragments (Weis et al., 1991; Sheriff et al., 1994; Weis & Drickamer, 1994), a human lung surfactant protein D fragment (Håkansson et al., 1999), a human E-selectin fragment (Graves et al., 1994), pancreatic stone protein (Bertrand et al., 1996), tetranectin (Nielsen et al., 1997), snake venom factor IX/X binding protein (Mizuno et al., 1997), tunicate C-type lectin (Poget et al., 1999), and from an NMR study of sea raven antifreeze protein (Gronwald et al., 1998). The fold of the C-type lectin domain is shown in Figure 4. The structure is mainly composed of two antiparallel β-sheets; one of these is four-stranded and the other five-stranded. The four-stranded β-sheet is found at the N-terminal part of the domain and is flanked by two helices. In the trimeric collectins, this sheet is interacting with the α-helical coiled-coil. The five-stranded β-sheet is rather distorted and more remotely located away from the trimer center. This β-sheet, together with some of the loop structure, makes up the carbohydrate binding ligand site. Two of the β-strands (β2 and β7) are relatively long and participate in both sheets. One of the disulfide bridges anchors helix α1 to strand β7 of the four-stranded β-sheet, the other ties together the two most peripheral strands (β5 and β6) of the five-stranded β-sheet. Since the calcium binding amino acid ligands are found at or close to these two strands, this disulfide bridge probably plays an important role in stabilizing the structure around the functional carbohydrate binding site.

A comparison of the amino acid sequences of 22 different extracellular collectins found in the SwissProt and Genbank databases and the recently sequenced pig SP-D (van Eijk et al., 2000) yields 18 conserved amino acids in addition to the four cysteines. These are indicated in Figure 4 using the numbering from human SP-D. Perhaps most important from a structural point of view is a hydrophobic cluster containing the conserved residues Phe304, Tyr314, Trp317, Pro322, Trp340, in addition to the semi-conserved Val332. They hold together the carbohydrate binding region consisting of strand β6 and the stretch between the two strands β4 and β5. Trp340 is also hydrogen bonded to the conserved asparagine residue 277 of helix α2. Asn277 is in turn hydrogen bonded to the main-chain carbonyl group of Gly338, as is the conserved residue Asn316. Glycine 338 appears to play a key role in the loop structure preceding the β6 strand. Its main-chain conformation is not compatible with the presence of a side chain, which is also true for

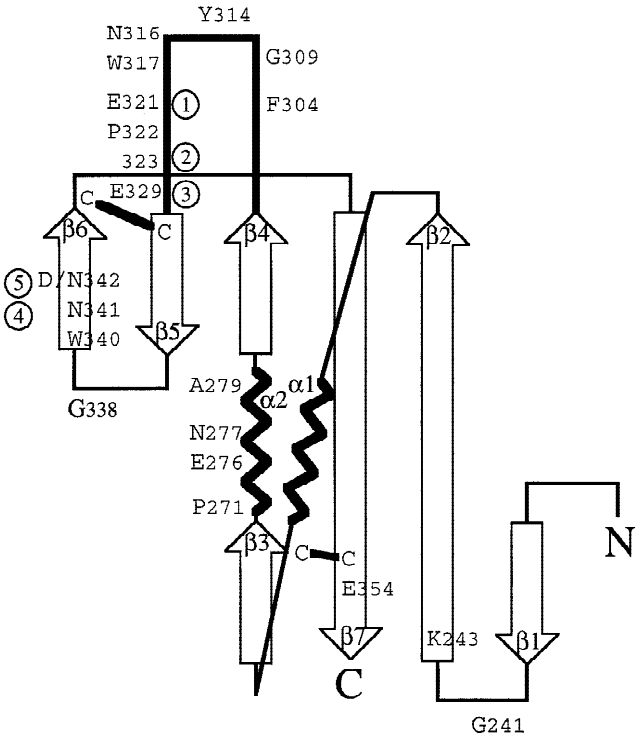

Fig. 4. Two-dimensional representation of the C-type lectin fold. The β-strands and α-helices have been numbered β1–β7 and α1–α2, respectively. Secondary structure elements were defined by PROCHECK (Laskowski et al., 1993) using the SP-D structure. They comprise the following residues (SP-D numbering); α1, 254–264; α2, 273–288; β1, 236–240; β2, 243–253; β3, 267–269; β4, 291–297; β5, 331–335; β6, 339-343; β7, 347–355. The amino acids that are coordinated to the carbohydrate binding calcium ion are denoted with encircled numbers 1–5. Twenty three different collectin sequences were aligned and conserved residues are indicated with their amino acid one letter symbol. The second calcium ligand is normally an asparagine but is an arginine or alanine in the SP-A sequences. The fifth calcium ligand is an aspartic acid residue in all sequences except dog SP-A, where it is an asparagine.

the other two conserved glycines, Gly241 and Gly309, and for most of the nonconserved SP-D glycines as well. Glu276, Asn277, and Ala279 are on the nonexposed side of helix α2, which is preceded by the conserved Pro271. Asn277 is, as already mentioned, hydrogen bonded to both Gly338 and Trp340, and Glu276 is hydrogen bonded to the main-chain amino group of the nonconserved Arg272 as well as to the side chain of Lys243. Lys243 forms salt bridges with both Glu276 and the likewise conserved Glu354. Thus, there is a hydrogen bond network of conserved residues linking β2 and β7 of the four-stranded β-sheet, helix α2 and β6 of the five-stranded β-sheet, as well as the important loop structures around β5 and β6. Surprisingly, the second of the calcium binding amino acid ligands is not conserved in SP-A, and the fifth ligand is an asparagine in canine SP-A but an aspartic acid in all other known collectin sequences. The fingerprint regions, discussed below, are found at the N-terminal side of β5 and the C-terminal end of β6, respectively. In CL-L1, Gly 241, Glu276, Ala279, Phe304, and Gly309 are all replaced, whereas all the consensus residues beyond Tyr314 are conserved even in this collectin. This indicates a slower evolutionary pace in and around the β5 and β6 strands, i.e., the carbohydrate binding site.

In addition to the carbohydrate binding calcium ion, two more calcium ions have been found in crystallographic studies of MBP

and SP-D (Weis et al., 1991; Weis & Drickamer, 1994; Håkansson et al., 1999). One of these is found ~8 Å away from the carbohydrate binding calcium and is complexed by three aspartic acid and one glutamic acid residues. These residues are either conserved or only conservatively replaced in the mannan-binding proteins, SP-D, conglutinin, and Cl-43, but not in SP-A. Two of the residues are also coordinated to the third calcium ion, but it is a matter of debate whether this calcium ion is present at physiological conditions or if it is an artifact of the high calcium ion concentration employed in the crystallographic analyses (Weis & Drickamer, 1994).

The electrostatic potential along the surface of these molecules displays a large positively charged area within the cavity between the three lectin domains. This surface charge can be observed both for the mannan-binding proteins and SP-D, but is not present on the C-type lectins that are not collectins, e.g., tetranectin or on the corresponding surface on the monomeric E-selectin (Fig. 5). In SP-D, this charge results from the presence of nonconserved lysine residues. The charged area may be involved in interactions with negative charges on the surface of its targets, e.g., the lipopolysaccharide moiety of microbial carbohydrates or the phospholipids of the pulmonary surfactant. The calcium ions in these proteins do influence the surface charge, and the apo protein is characterized by large negatively charged areas on the central and peripheral parts of the lectin domains (Fig. 5). Thus, the calcium ions might have a functional role other than carbohydrate binding, e.g., they may maintain a certain electrostatic potential pattern on the surface of the molecule. In this context, it is interesting to note that SP-D was found to bind to its putative receptor gp340 in a manner that was calcium dependent but not inhibited by the presence of maltose (Holmskov et al., 1997).

Interdomain contacts

There are very few intrachain contacts between the lectin domains and the coiled-coil. Each lectin domain contacts the α-helical coiled-coil through a hydrophobic cluster made up of residues from the β2 and β7 strands of the lectin domain and the coiled-coil region of an adjacent chain. None of the residues involved are conserved among the collectins but they are for the most part hydrophobic. A similar cluster made up of homologous residues is present in tetranectin, but in this case the lectin domain and coil segment involved are of the same chain and their interactions secured through a disulfide bond. Moreover, the same set of hydrophobic residues in the lectin domain mediate the formation of nonphysiological dimers in the crystallographic structure of MBP lectin domains lacking the coiled-coil region. The lectin domains are separated by a cavity on the threefold symmetry axis and do not contact each other. Thus, the interchain contacts are mediated through the N-terminal, collagenous and coiled-coil regions, but not the lectin domains.

Trimerization and oligomerization

The formation of a collagen triple helix is a slow process that requires seconds or even minutes. It is usually initiated by the association of an adjacent domain, which is C-terminal to the helix in the case of collagen (Engel & Prockop, 1991). In the collectins, this process is also initiated at the C-terminal end of the collagenous region through the formation of the α-helical coiled-coil. Formation of coiled-coils usually takes place within milliseconds and isolated SP-D neck regions have been shown to form stable trimers (Hoppe & Reid, 1994). Since the collagen triple helix is

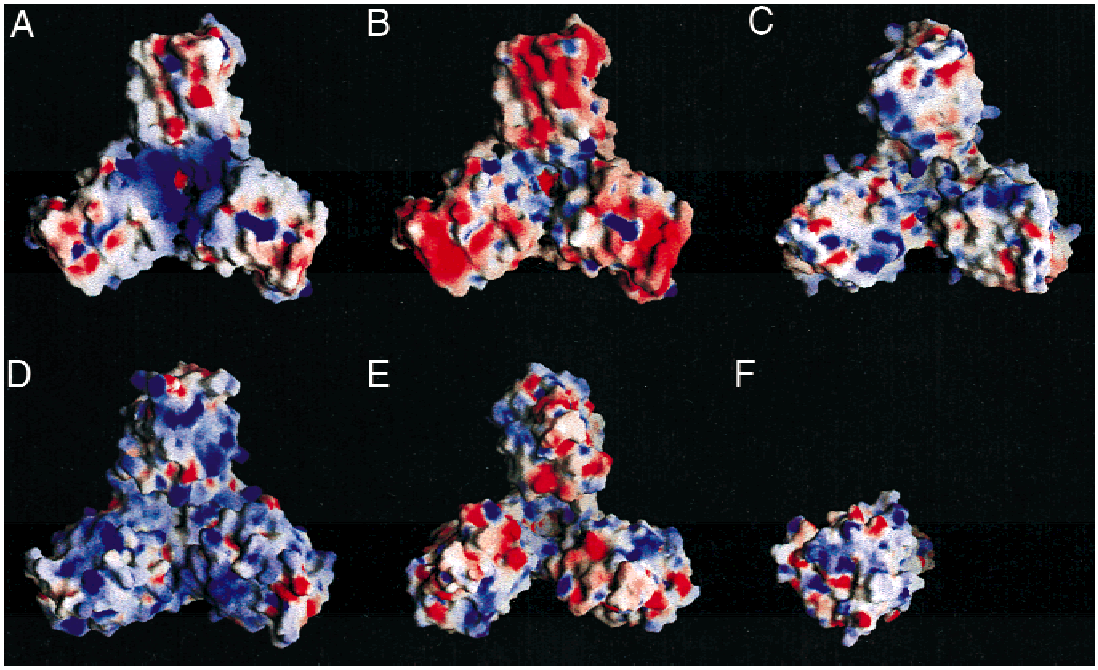

Fig. 5. Electrostatic surface potential of C-type lectins calculated by the program GRASP (Nicholls & Honig, 1991). Negative and positive potentials are shown in red and blue, respectively. **A:** SP-D with three calcium ions (per lectin domain). **B:** SP-D with no calcium ions. **C:** Human MBP with three calcium ions. **D:** Rat MBP-A with three calcium ions. **E:** Tetranectin with two calcium ions. **F:** E-selectin with three calcium ions. The trimeric molecules are viewed down the threefold symmetry axis, and the monomeric E-selectin from the same orientation.

not stable as a monomer, formation of helices and superhelices takes place simultaneously.

Isolated lectin domains do not trimerize; fragments of mannan-binding protein and tetranectin form trimers if and only if they contain at least a part of the coiled-coil region. Other C-type lectins lack a coiled-coil region and are not known to trimerize. The cavity between the lectin domains in the collagens is positively charged and suggests a repulsive interaction between the lectin domains (Fig. 5).

The assembly of oligomers of trimers is mediated by the N-terminal region and in the case of MBP-A also by the N-terminal part of the collagen triple helix. The final size and shape of the oligomers, which vary among the collectins, are also determined by the characteristics of their N-terminal regions. Oligomerization is an intracellular process; once secreted, the quaternary structure does not change (Wallis & Drickamer, 1999).

Carbohydrate binding site

The reason why these proteins are calcium dependent became evident as the structure of a complex between an MBP-A fragment and an oligosaccharide was determined (Weis et al., 1992). The positions of the carbohydrate binding sites in a mannan binding protein trimer are shown in Figure 6A, while Figure 6B offers a closer view of the binding site and the amino acid side chains involved. The binding site is situated in a shallow pocket peripherally located on the lectin domain some 30 Å away from the threefold symmetry axis (Sheriff et al., 1994; Ng et al., 1996;). The 3′ and 4′ hydroxyl groups of the terminal mannose residue of the carbohydrate are coordinated to a calcium ion bound to the protein. Interestingly, the space occupied by these hydroxyl groups, as well as the ether oxygen (O5), coincides with the positions of water molecules both in free mannan-binding protein and in SP-D. The calcium ion is bound to the lectin through five coordinated residues. In human SP-D, these residues are Glu321, Asn323, Glu329, Asn341, and Asp342. Since the numbering differs between different collectins, we shall refer to these residues and their homologues in other C-type lectins as ligands 1–5. In addition to these five side chains and the two water molecules or hydroxyl groups, the carbonyl oxygen of ligand 4 is also coordinated to the calcium ion. Ligands 1 and 2 are separated in the sequence only by a proline residue side chain embedded into a hydrophobic cluster

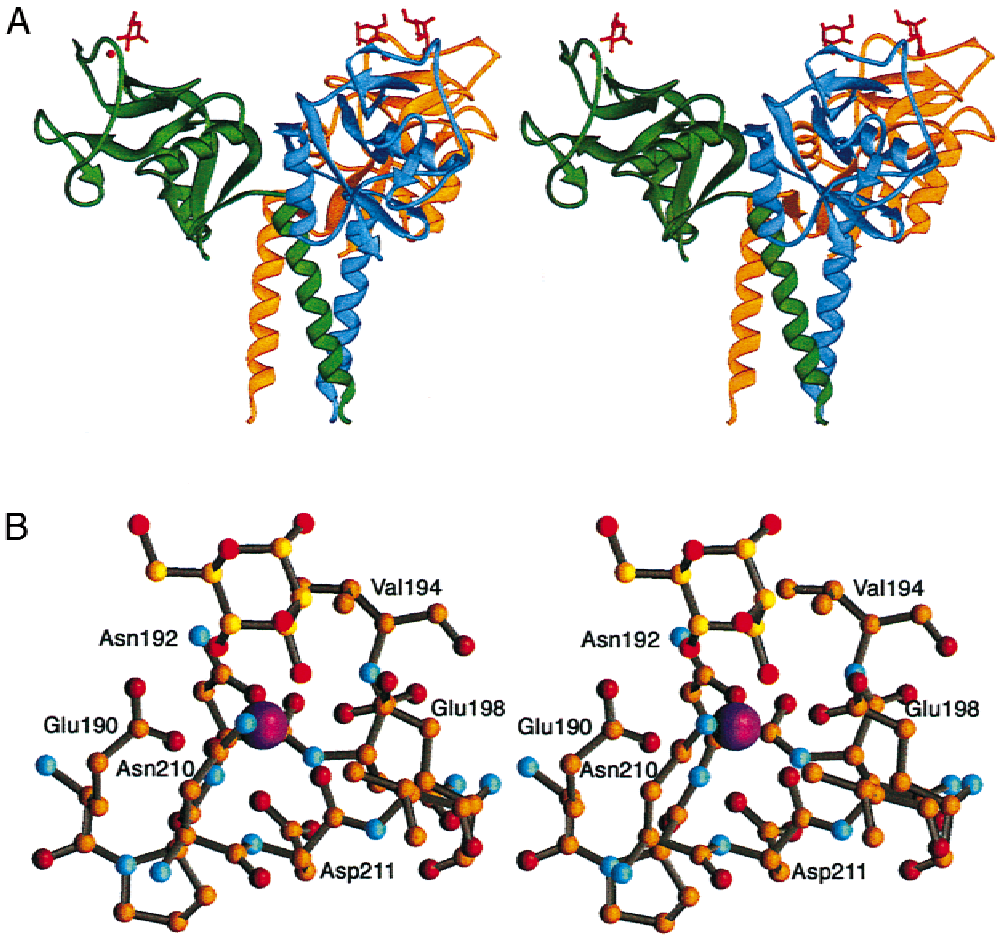

Fig. 6. A: The trimeric structure of a neck-CRD fragment of human MBP (Sheriff et al., 1994). Methyl-mannoside molecules, shown in red together with the calcium ions, have been modeled from the monomeric rat MBP-C complex (Ng et al., 1996). **B:** Methyl-mannoside bound to the carbohydrate binding site of rat MBP-C (Ng et al., 1996). The Glu190–Val194 segment, with its characteristic EPN motif (amino acid ligands 1 and 2) is shown behind the calcium ion. Asn210–Val212 (amino acid ligands 4 and 5) and Glu198 (amino acid ligand 3) are shown in front of and to the right of the calcium ion, respectively. The methyl aglycon of the carbohydrate ligand is not visible.

(*vide supra*). This motif presumably gives the coordination sphere a more stable framework, although its structure and the cis bond between ligand 1 and the proline are themselves stabilized by the calcium ion (Ng et al., 1998). Not all C-type lectins utilize all five of these ligands. For example, ligand 2 is replaced by an alanine in human SP-A and cannot therefore be a calcium ligand. Ligand 3 does not coordinate the calcium ion in selectin, although it is still a glutamate and ligand 5 is not a metal ligand in tetranectin where it is a phenylalanine.

The structures of a number of different MBP-carbohydrate complexes have since been elucidated. Collectin structure coordinates deposited with the Protein Data Bank are listed in Table 2. In all of the carbohydrate collectin complexes, two adjacent carbohydrate hydroxyl groups are coordinated to the calcium ion. Moreover, one of these hydroxyl groups is invariably within hydrogen bond distance to the side chains of ligands 1 (Glu) and 2 (Asn), while the other is within hydrogen bond distance to the side chains of ligands 3 (Glu), 4 (Asn), and to the main-chain carbonyl group of ligand 5. The relative positions of the carbohydrate ring differ between some of these structures, and different hydroxyl groups may be involved. As predicted (Weis et al., 1992), the common denominator is ligation of two vicinal, equatorial hydroxyl groups in the (+)-synclinical configuration that replaces two water molecules bound to the 8-coordinated calcium ion.

In addition to the coordination of the two hydroxyl groups to the calcium ion, there appears to be few contacts between carbohydrate and protein. The only other general interactions between the carbohydrate ring and MBP appear to be van der Waal contacts involving His189 and Ile207 in MBP-A (Weis et al., 1992), and Val194 in MBP-C (Ng et al., 1996). The homologous His189 in MBP-A and Val194 in MBP-C superimpose well with the corresponding aspartic acid residue in human SP-D (Håkansson et al., 1999). This residue has no counterpart in SP-A, but conglutinin and CL-43 both contain a small insertion in this part of the sequence. In galactose specific C-type lectins, this residue is replaced by a tryptophan and hosts a glycine rich loop insertion, as discussed below. His189 in MBP-A and the following residues have been referred to as a fingerprint motif that distinguishes the different collectins from each other (Håkansson et al., 1999). The position corresponding to Ile207 in MBP-A marks the beginning of a second fingerprint motif. In SP-D, this is an arginine that, with its more extended and polar side chain, has the potential of interacting with residues beyond the terminal mannose residue. It is possible that these fingerprint regions modulate and fine-tune the carbohydrate specificities throughout the collectin family. On the other hand, the only observed carbohydrate–protein interaction involving a carbohydrate residue other than the terminal mannose is a hydrogen bond between an ill-defined lysine residue in MBP-A and the penultimate residue in high-mannose oligosaccharide. All other interactions between protein and nonterminal mannose residues are mediated by water molecules and are therefore not very specific (Weis et al., 1992). The paucity of atomic interactions between carbohydrate and protein is also in line with the generally weak and nonspecific binding of carbohydrate to a collectin monomer. The structure of a selectin-like mutant (Ng & Weis, 1997) complexed to sialyl-Lewis^x oligosaccharide has shown that a ly-

Table 2. *Collectin structure coordinates deposited with the Protein Data Bank (Bernstein et al., 1977)*[a]

Protein	Trimer	Species	Ligand[a]	PDB code	Reference
MBP-A	−	rat	Ho(III) substituted	1MSB	Weis et al. (1991)
MBP-A	−	rat	Oligosaccharide	2MSB	Weis et al. (1992)
MBP	+	human		1HUP	Sheriff et al. (1994)
MBP-C	−	rat	α-Methyl-L-fucopyranoside	1RDI	Ng et al. (1996)
MBP-C	−	rat	β-Methyl-L-fucopyranoside	1RDJ	Ng et al. (1996)
MBP-C	−	rat	D-Galactose	1RDK	Ng et al. (1996)
MBP-C	−	rat	α-Methyl-D-mannopyranoside	1RDL	Ng et al. (1996)
MBP-C	−	rat	α-Methyl-D-mannopyranoside	1RDM	Ng et al. (1996)
MBP-C	−	rat	α-Methyl-D-N-acetylglucosaminide	1RDN	Ng et al. (1996)
MBP-C	−	rat		1RDO	Ng et al. (1996)
MBP-A	+	rat		1RTM	Weis and Drickamer (1994)
MBP-A	−	rat	Yt(III) substituted	1YTT	Burling et al. (1996)
MBP-A	+	rat	Ho(III) substituted	1BUU	Ng et al. (1998)
MBP-C	−	rat	Apo	1BV4	Ng et al. (1998)
MBP-A	+	rat mutant[a]	β-Methyl-D-galactoside	1AFA	Kolatkar and Weis (1996)
MBP-A	+	rat mutant[a]	N-Acetyl-D-galactosamine	1AFB	Kolatkar and Weis (1996)
MBP-A	+	rat mutant[a]		1AFD	Kolatkar and Weis (1996)
MBP-A	+	rat mutant[b]	N-Acetyl-D-galactosamine	1BCH	Kolatkar et al. (1998)
MBP-A	+	rat mutant[c]	N-Acetyl-D-galactosamine	1BCJ	Kolatkar et al. (1998)
MBP-A	+	rat mutant[d]		1KMB	Ng and Weis (1997)
MBP-A	+	rat mutant[d]	3′Neuac-Lewis-X	2KMB	Ng and Weis (1997)
MBP-A	+	rat mutant[d]	3′-Sulfo-Lewis-X	3KMB	Ng and Weis (1997)
MBP-A	+	rat mutant[d]	4′-Sulfo-Lewis-X	4KMB	Ng and Weis (1997)
SP-D	+	human		1B08	Håkansson et al. (1999)

[a] "Galactose-specific mutant," E185Q, N187D, H189W, G190Y, S191G, Ins(H192, G193, L194, G195, G196).
[b] Mutant a with T202H.
[c] Mutant a with S154V, T202H.
[d] "Selectin-like mutant," A211K, S212K, H213K. "Substituted" and "Apo" refer to replacement and removal of native calcium ion.

sine residue belonging to the second fingerprint region does indeed interact with the oligosaccharide. The more charged and hydrophilic fingerprint regions of SP-D also present a surface with more hydrogen bonding opportunities compared to MBP, but no SP-D–ligand complex has yet been visualized. At this point we have to conclude that less specific water-mediated contacts account for most if not all interactions between protein and nonterminal carbohydrate residues.

Carbohydrate binding and specificity

Methyl-mannose binds to mannan-binding protein C with an orientation that is different from that of an oligosaccharide terminal mannose group bound to mannan-binding protein A. In both cases, the carbohydrate interacts with the calcium ion with the 3′ and 4′ hydroxyl groups but their relative orientations are reversed, i.e., the 4′ hydroxyl group of methyl-mannose occupies the same position as the 3′ hydroxyl group of the oligosaccharide complexed to MBP-A and vice versa. Although the mannose residue rings are superimposable, their relative orientations differ by ~180°. This structural diversity is possible because of the absence of additional close interactions between carbohydrate and protein. The difference between these two ligand structures has been interpreted as due to a difference between MBP-A and MBP-C. However, the carbohydrate moieties and the experimental conditions also differ between these two studies. In particular, the presence of His189 in MBP-A might interfere with the α1-anomeric linkage if the orientation of the oligosaccharide is changed. This does not, however, prevent the sialyl-Lewis^x oligosaccharide from binding to a selectin-like mutant (Ng & Weis, 1997) with the α1-anomeric carbon directed toward His189. An alternative interpretation is that there are multiple ways for carbohydrates to bind, with coordination of two adjacent equatorial hydroxyl groups as the only common denominator. In solution, carbohydrates can move freely and bind in the chemically most favorable way. However, the natural carbohydrate ligands for collectins are linked to a target surface, and the chemically most favored binding might, due to this restriction, not always be possible for all carbohydrate binding domains that are involved. In fact, the high mannose oligosaccharide used in the experiment is bridging two crystallographically nonequivalent lectin domains, and the observed ligand orientation might be influenced by this linkage. The existence of alternative ways of binding would enable the collectins to bind two or three lectin domains to a target surface lacking threefold symmetry. Moreover, if one way of binding is prohibited by steric exclusion between protein and nonterminal residues, a different way of binding could be adopted. The lack of specific interactions between protein and nonterminal oligosaccharide residues strengthens this idea.

The collectins are rather similar in their carbohydrate specificity. In particular, they bind more strongly to mannose residues than to galactose residues. However, the reverse is true for some other C-type lectins, e.g., the hepatocytic asialoglycoprotein receptor. It was shown in an elegant study that mutating the amino acid ligands 1 and 2 within the EPN motif in mannan-binding protein to QPD, as found in the asialoglycoprotein receptor, resulted in a protein that preferred galactose over mannose (Drickamer, 1992). Additional mutations, i.e., H189W and the insertion of a glycine rich loop (see footnotes to Table 2), yielded an affinity comparable to the receptor protein (Iobst & Drickamer, 1994). The structures of a complex between this mutant and two galactose derivatives, methyl-D-galactoside and N-acetyl-D-galactosamine have been analyzed crystallographically (Kolatkar & Weis, 1996). These derivatives bind to the mutant with their 3′ and 4′ hydroxyl groups coordinated to the carbon in a manner similar to mannose and the wild-type protein. These hydroxyl groups are both equatorial in mannose but equatorial and axial, respectively, in galactose. As a result, the overall position and orientation of the sugar ring of galactose is very different from that of mannose. The presence of the large indole ring of the H189W mutation was found to discriminate against mannose but interact favorably with the apolar face of the galactose residue. The effect of the reversal of the acid-amide pattern of the amino acid ligands 1 and 2 from EPN to QPD is more difficult to explain. While inspecting these structures, we noticed that the geometry around the coordinated hydroxyl group is better (i.e., more tetrahedral) for the hydrogen bond acceptor (the acid) than for the donor (the amide) in both the wild-type mannoside complex and the mutant galactoside complex. The nucleophilicity of the neutral hydroxyl group is probably increased by the presence of a strong hydrogen bond involving its proton. In either case, the acid-amide pattern seems to enhance the nucleophilicity, and hence calcium binding of this hydroxyl group. Amino acid ligand 1 is strictly conserved in the collectin group of proteins. Although mutation of the amino acid ligand 2 from an asparagine to an aspartic acid results in a loss of affinity for mannose (Quesenberry & Drickamer, 1992), this residue is an alanine in human SP-A and an arginine most other SP-A sequences. Poget et al. (1999) offer a another explanation for the monosaccharide specificities, namely that the hydrogen of the sugar hydroxyl is always oriented away from the ring. In their structure of tunicate C-type lectin complexed with galactose, the hydrogen bond geometry is again better for the bond where the hydroxyl group act as a donor rather than acceptor, although this lectin has an EPSN sequence rather than QPD.

Deviation from threefold symmetry

The 3D structure of a lung surfactant protein D trimeric fragment consisting of the coiled-coil region and lectin domains displays some interesting deviations from symmetry (Håkansson et al., 1999). The side-chain conformation of Tyr228 of the α-helical coiled-coil differs between the three peptide chains. This residue is in a "d" position in the heptad repeat. The side chain of one of the three tyrosines is completely buried in the interior of the coil, while the other two are partly exposed to the solvent. The partly exposed tyrosines are able to form hydrogen bonds with their hydroxyl groups. They are, unlike the buried tyrosine, within van der Waal's distance to Ile244 of the lectin domain. These two lectin domains, due to the steric interference of the Tyr228–Ile244 contact, are not as close to the coiled-coil as the third lectin domain. Does this represent a case of static deviation from threefold symmetry or are the lectin domains allowed to move within a relative narrow range, restricted but not fixed by the tyrosine side chain? Can the symmetry violation be influenced by ligand binding to a target surface? Can it help the collectins to better match a target carbohydrate array? Can it be propagated further to the asymmetric junction between coiled-coil and triple helix? These questions cannot be resolved until more structural data of SP-D and the other collectins become available.

References

Beck K, Brodsky B. 1998. Supercoiled protein motifs: The collagen triple-helix and the alpha-helical coiled coil. *J Struct Biol 122*:17–29.

Bella J, Eaton M, Brodsky B, Berman HM. 1994. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution [see Comments]. *Science 266*:75–81.

Bernstein FC, Koetzle TF, Williams GJ, Meyer EE Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bertrand JA, Pignol D, Bernard JP, Verdier JM, Dagorn JC, Fontecilla-Camps JC. 1996. Crystal structure of human lithostathine, the pancreatic inhibitor of stone formation. *EMBO J 15*:2678–2684.

Burling FT, Weis WI, Flaherty KM, Brünger AT. 1996. Direct observation of protein solvation and discrete disorder with experimental crystallographic phases. *Science 271*:72–77.

Colley KJ, Baenziger JU. 1987. Identification of the post-translational modifications of the core-specific lectin. The core-specific lectin contains hydroxyproline, hydroxylysine, and glucosylgalactosylhydroxylysine residues. *J Biol Chem 262*:10290–10295.

Crouch E, Persson A, Chang D, Heuser J. 1994. Molecular structure of pulmonary surfactant protein D (SP-D). *J Biol Chem 269*:17311–17319.

Crouch EC. 1998. Collectins and pulmonary host defense. *Am J Respir Cell Mol Biol 19*:177–201.

Drickamer K. 1992. Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature 360*:183–186.

Elhalwagi BM, Damodarasamy M, McCormack FX. 1997. Alternate amino terminal processing of surfactant protein A results in cysteinyl isoforms required for multimer formation. *Biochemistry 36*:7018–7025.

Engel J, Prockop DJ. 1991. The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. *Annu Rev Biophys Biophys Chem 20*:137–152.

Graves BJ, Crowther RL, Chandran C, Rumberger JM, Li S, Huang KS, Presky DH, Familletti PC, Wolitzky BA, Burns DK. 1994. Insight into E-selectin/ligand interaction from the crystal structure and mutagenesis of the lec/EGF domains. *Nature 367*:532–538.

Gronwald W, Loewen MC, Lix B, Daugulis AJ, Sonnichsen FD, Davies PL, Sykes BD. 1998. The solution structure of type II antifreeze protein reveals a new member of the lectin family. *Biochemistry 37*:4712–4721.

Håkansson K, Lim NK, Hoppe HJ, Reid KB. 1999. Crystal structure of the trimeric alpha-helical coiled-coil and the three lectin domains of human lung surfactant protein D. *Struct Fold Des 7*:255–264.

Holmskov U, Laursen SB, Malhotra R, Wiedemann H, Timpl R, Stuart GR, Tornoe I, Madsen PS, Reid KB, Jensenius JC. 1995. Comparative study of the structural and functional properties of a bovine plasma C-type lectin, collectin-43, with other collectins. *Biochem J 305*:889–896.

Holmskov U, Lawson P, Teisner B, Tornoe I, Willis AC, Morgan C, Koch C, Reid KB. 1997. Isolation and characterization of a new member of the scavenger receptor superfamily, glycoprotein-340 (gp-340), as a lung surfactant protein-D binding molecule. *J Biol Chem 272*:13743–13749.

Hoppe HJ, Reid KB. 1994. Trimeric C-type lectin domains in host defence. *Structure 2*:1129–1133.

Iobst ST, Drickamer K. 1994. Binding of sugar ligands to Ca(2+)-dependent animal lectins. II. Generation of high-affinity galactose binding by site-directed mutagenesis. *J Biol Chem 269*:15512–15519.

Kolatkar AR, Leung AK, Isecke R, Brossmer R, Drickamer K, Weis WI. 1998. Mechanism of N-acetylgalactosamine binding to a C-type animal lectin carbohydrate-recognition domain. *J Biol Chem 273*:19502–19508.

Kolatkar AR, Weis WI. 1996. Structural basis of galactose recognition by C-type animal lectins. *J Biol Chem 271*:6679–6685.

Kramer RZ, Bella J, Mayville P, Brodsky B, Berman HM. 1999. Sequence dependent conformational variations of collagen triple-helical structure. *Nat Struct Biol 6*:454–457.

Kramer RZ, Vitagliano L, Bella J, Berisio R, Mazzarella L, Brodsky B, Zagari A, Berman HM. 1998. X-ray crystallographic determination of a collagen-like peptide with the repeating sequence (Pro-Pro-Gly). *J Mol Biol 280*:623–638.

Kurata H, Sannoh T, Kozutsumi Y, Yokota Y, Kawasaki T. 1994. Structure and function of mannan-binding proteins isolated from human liver and serum. *J Biochem (Tokyo) 115*:1148–1154.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Lee YM, Leiby KR, Allar J, Paris K, Lerch B, Okarma TB. 1991. Primary structure of bovine conglutinin, a member of the C-type animal lectin family. *J Biol Chem 266*:2715–2723.

Lu J. 1997. Collectins: collectors of microorganisms for the innate immune system. *Bioessays 19*:509–518.

Lu J, Wiedemann H, Timpl R, Reid KB. 1993. Similarity in structure between C1q and the collectins as judged by electron microscopy. *Behring Inst Mitt 93*:6–16.

Lu JH, Thiel S, Wiedemann H, Timpl R, Reid KB. 1990. Binding of the pentamer/hexamer forms of mannan-binding protein to zymosan activates the proenzyme C1r2C1s2 complex, of the classical pathway of complement, without involvement of C1q. *J Immunol 144*:2287–2294.

Malhotra R, Laursen SB, Willis AC, Sim RB. 1993. Localization of the receptor-binding site in the collectin family of proteins. *Biochem J 293*:15–19.

Matsushita M, Fujita T. 1992. Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. *J Exp Med 176*:1497–1502.

McCormack FX, Calvert HM, Watson PA, Smith DL, Mason RJ, Voelker DR. 1994. The structure and function of surfactant protein A. Hydroxyproline- and carbohydrate-deficient mutant proteins. *J Biol Chem 269*:5833–5841.

McCormack FX, Damodarasamy M, Elhalwagi BM. 1999. Deletion mapping of N-terminal domains of surfactant protein A. The N-terminal segment is required for phospholipid aggregation and specific inhibition of surfactant secretion. *J Biol Chem 274*:3173–3181.

McCormack FX, Pattanajitvilai S, Stewart J, Possmayer F, Inchley K, Voelker DR. 1997. The Cys6 intermolecular disulfide bond and the collagen-like region of rat SP-A play critical roles in interactions with alveolar type II cells and surfactant lipids. *J Biol Chem 272*:27971–27979.

Mizuno H, Fujimoto Z, Koizumi M, Kano H, Atoda H, Morita T. 1997. Structure of coagulation factors IX/X-binding protein, a heterodimer of C-type lectin domains [Letter]. *Nat Struct Biol 4*:438–441.

Ng KK, Drickamer K, Weis WI. 1996. Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. *J Biol Chem 271*:663–674.

Ng KK, Park-Snyder S, Weis WI. 1998. Ca2+-dependent structural changes in C-type mannose-binding proteins. *Biochemistry 37*:17965–17976.

Ng KK, Weis WI. 1997. Structure of a selectin-like mutant of mannose-binding protein complexed with sialylated and sulfated Lewis(x) oligosaccharides. *Biochemistry 36*:979–988.

Nicholls A, Honig BJ. 1991. A rapid finite-difference algoritim, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. *J Comput Chem 12*:435–445.

Nielsen BB, Kastrup JS, Rasmussen H, Holtet TL, Graversen JH, Etzerodt M, Thogersen HC, Larsen IK. 1997. Crystal structure of tetranectin, a trimeric plasminogen-binding protein with an alpha-helical coiled coil. *FEBS Lett 412*:388–396.

Ohtani K, Suzuki Y, Eda S, Kawai T, Kase T, Yamazaki H, Shimada T, Keshi H, Sakai Y, Fukuoh A, et al. 1999. Molecular cloning of a novel human collectin from liver (CL-L1). *J Biol Chem 274*:13681–13689.

Persson A, Chang D, Rust K, Moxley M, Longmore W, Crouch E. 1989. Purification and biochemical characterization of CP4 (SP-D), a collagenous surfactant-associated protein. *Biochemistry 28*:6361–6367.

Poget SF, Legge GB, Proctor MR, Butler PJ, Bycroft M, Williams RL. 1999. The structure of a tunicate C-type lectin from Polyandrocarpa misakiensis complexed with D-galactose. *J Mol Biol 290*:867–879.

Quesenberry MS, Drickamer K. 1992. Role of conserved and nonconserved residues in the Ca(2+)-dependent carbohydrate-recognition domain of a rat mannose-binding protein. Analysis by random cassette mutagenesis. *J Biol Chem 267*:10831–10841.

Rothmann AB, Mortensen HD, Holmskov U, Hojrup P. 1997. Structural characterization of bovine collectin-43. *Eur J Biochem 243*:630–635.

Sano K, Fisher J, Mason RJ, Kuroki Y, Schilling J, Benson B, Voelker D. 1987. Isolation and sequence of a cDNA clone for the rat pulmonary surfactant-associated protein (PSP-A). *Biochem Biophys Res Commun 144*:367–374.

Sheriff S, Chang CY, Ezekowitz RA. 1994. Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil. *Nat Struct Biol 1*:789–794.

Thiel S, Vorup-Jensen T, Stover CM, Schwaeble W, Laursen SB, Poulsen K, Willis AC, Eggleton P, Hansen S, Holmskov U, et al. 1997. A second serine protease associated with mannan-binding lectin that activates complement. *Nature 386*:506–510.

van Eijk M, Haagsman HP, Skinner T, Archibold A, Reid KB, Lawson PR. 2000. Porcine lung surfactant protein D: Complementary DNA cloning, chromosomal localization, and tissue distribution. *J Immunol 164*:1442–1450.

Voss T, Eistetter H, Schafer KP, Engel J. 1988. Macromolecular organization of natural and recombinant lung surfactant protein SP 28-36. Structural homology with the complement factor C1q. *J Mol Biol 201*:219–227.

Wallis R, Drickamer K. 1997. Asymmetry adjacent to the collagen-like domain in rat liver mannose-binding protein. *Biochem J 325*:391–400.

Wallis R, Drickamer K. 1999. Molecular determinants of oligomer formation and complement fixation in mannose-binding proteins. *J Biol Chem 274*:3580–3589.

Weis WI, Drickamer K. 1994. Trimeric structure of a C-type mannose-binding protein. *Structure 2*:1227–1240.

Weis WI, Drickamer K. 1996. Structural basis of lectin-carbohydrate recognition. *Annu Rev Biochem 65*:441–473.

Weis WI, Drickamer K, Hendrickson WA. 1992. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature 360*:127–134.

Weis WI, Kahn R, Fourme R, Drickamer K, Hendrickson WA. 1991. Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science 254*:1608–1615.

Zhang M, Damodarasamy M, Elhalwagi BM, McCormack FX. 1998. The longer isoform and Cys-1 disulfide bridge of rat surfactant protein A are not essential for phospholipid and type II cell interactions. *Biochemistry 37*:16481–16488.