The EMBO Journal (2007) 26, 623–633 | © 2007 European Molecular Biology Organization | All Rights Reserved 0261-4189/07
www.embojournal.org

THE EMBO JOURNAL

Structural insights into the innate immune recognition specificities of L- and H-ficolins

Virginie Garlatti[1], Nicolas Belloy[1], Lydie Martin[1], Monique Lacroix[2], Misao Matsushita[3], Yuichi Endo[4], Teizo Fujita[4], Juan Carlos Fontecilla-Camps[1], Gérard J Arlaud[2], Nicole M Thielens[2] and Christine Gaboriaud[1,*]

[1]Laboratoire de Cristallographie et Cristallogénèse des Protéines, Grenoble, France, [2]Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, CEA; CNRS; Université Joseph Fourier, Grenoble, France, [3]Department of Applied Biochemistry, Institute of Glycotechnology, Tokai University, Hiratsuka, Kanagawa, Japan and [4]Department of Biochemistry, Fukushima Medical University School of Medicine, Fukushima, Japan

Innate immunity relies critically upon the ability of a few pattern recognition molecules to sense molecular markers on pathogens, but little is known about these interactions at the atomic level. Human L- and H-ficolins are soluble oligomeric defence proteins with lectin-like activity, assembled from collagen fibers prolonged by fibrinogen-like recognition domains. The X-ray structures of their trimeric recognition domains, alone and in complex with various ligands, have been solved to resolutions up to 1.95 and 1.7 Å, respectively. Both domains have three-lobed structures with clefts separating the distal parts of the protomers. Ca^{2+} ions are found at sites homologous to those described for tachylectin 5A (TL5A), an invertebrate lectin. Outer binding sites (S1) homologous to the GlcNAc-binding pocket of TL5A are present in the ficolins but show different structures and specificities. In L-ficolin, three additional binding sites (S2–S4) surround the cleft. Together, they define an unpredicted continuous recognition surface able to sense various acetylated and neutral carbohydrate markers in the context of extended polysaccharides such as 1,3-β-D-glucan, as found on microbial or apoptotic surfaces.

The EMBO Journal (2007) **26,** 623–633. doi:10.1038/sj.emboj.7601500; Published online 11 January 2007
Subject Categories: immunology; structural biology
Keywords: carbohydrate recognition; fibrinogen-like domain; innate immunity; N-acetyl group; X-ray crystallography

Introduction

The ability to discriminate infectious agents from self is intrinsic to innate immunity (Janeway, 1992). This crucial task is fulfilled by pattern recognition molecules that sense invariant structures exposed at microbial surfaces, such as lipopolysaccharides, peptidoglycans, lipoteichoic acids, mannans, and 1,3-β-D-glucans. Recognition elicits effector mechanisms designed to provide collectively a frontline defence against infection (Hoffmann *et al*, 1999). With a few exceptions such as peptidoglycan and lipopolysaccharide recognition proteins (Guan *et al*, 2005; Kim *et al*, 2005) and various C-type lectins (mannan-binding lectin (MBL) (Takahashi *et al*, 2005), lung surfactant proteins (Shrive *et al*, 2003), and dectin 1 (Adachi *et al*, 2004)), precise structural information about the mechanisms that allow a restricted number of recognition proteins to sense these danger signals is lacking. For all recognition molecules investigated so far, ligand binding involves highly specific interactions. One of the best documented cases is that of MBL (Weis *et al*, 1992), which through its C-type lectin domain recognizes terminal carbohydrates, provided that their hydroxyl groups at positions C3 and C4 are in the equatorial orientation.

One of the major routes of innate host defence in humans, the lectin pathway of complement activation, is triggered through the action of MBL-associated serine protease-2 in response to recognition of neutral carbohydrates and other motifs present on microbial surfaces (Thiel *et al*, 1997; Fujita, 2002; Holmskov *et al*, 2003). Four different pattern recognition molecules initiate this pathway: MBL, L-ficolin, H-ficolin, and M-ficolin (Matsushita and Fujita, 1992; Matsushita *et al*, 2000, 2002; Liu *et al*, 2005; Frederiksen *et al*, 2005). MBL is an oligomer of trimers, with each subunit consisting of a collagen-like domain, a 'neck' region, and a carbohydrate recognition domain of the C-lectin type (Turner, 1996). Ficolins exhibit a similar oligomeric structure, but their collagen-like stalks are followed by a domain homologous to the fibrinogen β and γ chains (Matsushita and Fujita, 2001). L-ficolin/P35 (or ficolin 2) and H-ficolin (or ficolin 3) are serum proteins, whereas M-ficolin (or ficolin 1) is a secretory protein produced by lung and blood cells (Matsushita *et al*, 1996; Sugimoto *et al*, 1998; Liu *et al*, 2005).

L-ficolin was shown to activate the lectin pathway after binding to various capsulated bacteria (Matsushita *et al*, 1996; Lynch *et al*, 2004; Aoyagi *et al*, 2005). A binding specificity for GlcNAc was initially characterized (Matsushita *et al*, 1996, 2000; Le *et al*, 1998), in agreement with the finding that L-ficolin recognizes lipoteichoic acid, a GlcNAc-containing cell wall component characteristic of Gram-positive bacteria (Lynch *et al*, 2004). Binding to a fungal 1,3-β-D-glucan preparation was later reported (Ma *et al*, 2004), whereas further investigations have revealed specificity for N-acetylated carbohydrates and other non-carbohydrate acetylated compounds such as acetylcholine (Krarup *et al*, 2004).

M-ficolin is highly homologous to L-ficolin at the amino-acid sequence level and also shows a marked preference for acetylated compounds (Frederiksen *et al*, 2005; Liu *et al*, 2005). H-ficolin, in contrast, is less closely related to L-ficolin and shows no binding affinity for acetylated derivatives. So

*Corresponding author. Laboratoire de Cristallographie et Cristallogénèse des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel, 41, rue Jules Horowitz, Grenoble cedex 1, 38041, France. Tel.: +33 4 38 78 95 99; Fax: +33 4 38 78 51 22; E-mail: christine.gaboriaud@ibs.fr

Received: 2 August 2006; accepted: 31 October 2006; published online: 11 January 2007

far, it has been only reported to bind to *Aerococcus viridans* (Tsujimura *et al*, 2001; Krarup *et al*, 2005).

The above observations indicate that ficolins are not classical lectins and raise questions about their mechanism of action. To answer these questions at the molecular level, we have produced the recombinant recognition domains of L- and H-ficolins and solved their structures by X-ray crystallography, alone and in complex with different ligands. Both proteins are homotrimers of protomers homologous to tachylectin 5A (TL5A), a fibrinogen-like lectin from the invertebrate *Tachypleus tridentatus* (Kairies *et al*, 2001). The structures reveal that L-ficolin has evolved to be a versatile recognition protein able to recognize a variety of acetylated and carbohydrate targets through three different sites, suggesting that ficolins represent novel types of pattern recognition proteins.

Results and discussion

We have determined the three-dimensional structure of the fibrinogen-like recognition domains of human L- and H-ficolins. The segments corresponding to these domains (residues 70–288 and 58–276 of mature L- and H-ficolins, respectively) were expressed in a baculovirus/insect cells system. A single H-ficolin species was produced, with a mass ranging from 25 932 to 26 231 ± 12 Da, as determined by mass spectrometry, accounting for the polypeptide chain (calculated mass: 25 078 Da) plus a heterogeneous N-linked oligosaccharide (854–1153 ± 12 Da). Two different species were recovered in the case of L-ficolin, a minor species with a mass of 24 906 ± 12 Da corresponding to the unglycosylated form (calculated mass: 24 916.6 Da), and a major one resolved into four peaks ranging from 25 805 to 26 311 ± 12 Da, consistent with the presence of a heterogeneous oligosaccharide with a mass of 888–1394 ± 12 Da. Analysis of the glycosylated L-ficolin species by sedimentation velocity yielded a sedimentation coefficient ($s_{20,w}$) of 5.1 and a molecular mass calculated from s and D (diffusion coefficient) of 68 ± 7 kDa, consistent with a trimer.

Overall structures

Four different crystal forms were generated using the native, glycosylated L-ficolin fragment alone or in the presence of various ligands, and one form was obtained for H-ficolin (Table I). The L-ficolin structure was initially solved by molecular replacement using TL5A (Kairies *et al*, 2001) as a search model and subsequently refined to a resolution up to 1.95 Å. The H-ficolin structure was then solved using L-ficolin as a model and refined at 1.7 Å resolution (see Materials and methods). Crystallographic data collection and refinement

Table I Characteristics of the crystals used in this study[a]

Ficolin crystal (form)	Resolution (Å)	R_{work} R_{free} (%)	Ligand[b] concentration (mM)	Site	PDB code[c]
H-1[d]	2.2	20.1 26.5	None	None	2j64
H-2[d]	2.35	21.2 25.8	D-Fucose 50 (C)	S1	2j60
H-3[d]	1.75	22.2 25.9	L-Fucose 50 (C)	S1[e]	
H-4[d]	1.7	20.1 23.2	Galactose 25 (C)	S1	2j5z
L-1 (A)[f]	2.5	20.2 25.9	None	S1[g]	2j3g
L-2 (A)[f]	1.95	17.6 21.0	CysNAc 0.25 (C)	S2	2j1g
L-3 (A)[f]	2.15	18.8 23.7	Galactose 150 (S)	S2 (S3, S4)	2j3u
L-4 (A)[f]	2.35	20.8 24.8	β-Glucan 150 (S)	S3–S4	2j0y
L-5 (A)[f]	2.85	19.6 24.8	ManNAc 10 (S)	S1[g] S3	2j0g
L-6 (A)[f]	2.85	20.8 27.4	Acetylcholine 1 (C)	S3	2j0h
L-7 (A)[f]	2.8	22.1 27.6	GalNAc 50 (S)	S3	2j3f
L-8 (B)[h]	2.65	22.7 28.2	GlcNAc 10 (C)	S2	2j3o
L-9 (C)[i]	2.7	20.6 29.1	GlcNAc 10 (C)	S2	2j61
L-10 (A)[f]	2.8	19.7 26.9	CysNAc 150 (S)	S1[g], S2, S3, others	2j2p

[a]The nature of the ligand, its concentration, the protocol used, and the site location of the ligand are detailed, as well as the resolution and R_{work}/R_{free} refinement statistics. Complete crystallographic statistics are provided in Supplementary Table II.
[b]Co-crystallization (C) or soaking (S) experiments.
[c]ID code in the Protein Data Bank (www.rcsb.org).
[d]H-ficolin crystals: $P4_12_12$ $a=b=62$ Å, $c=347$ Å, $\alpha=\beta=\gamma=90^\circ$.
[e]Glycerol (used for cryoprotection).
[f]L-ficolin crystal form A: $P3_2$, $a=b=99$ Å, $c=142$ Å, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$.
[g]Mannose from a neighboring molecule in the crystal (see text and Figure 3C).
[h]L-ficolin crystal form B: C2 $a=200.9$ Å, $b=84.3$ Å, $c=144.9$ Å, $\alpha=90^\circ$, $\beta=124.5^\circ$, $\gamma=90^\circ$.
[i]L-ficolin crystal form C: $P6_3$ $a=79.5$ Å, $b=79.5$ Å, $c=172$ Å, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$.

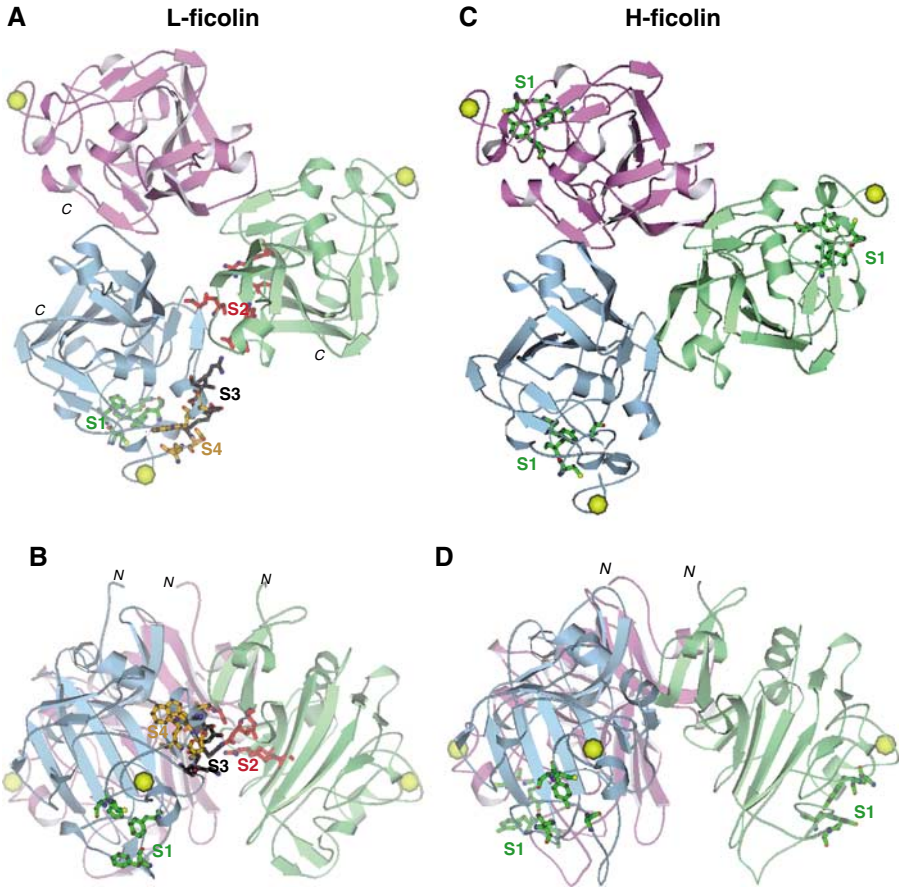

Figure 1 Homotrimeric structure of the recognition domains of human L- and H-ficolins and location of their binding site(s). (**A, B**) L-ficolin structure seen from the target binding surface (bottom view) and on a perpendicular side view. (**C, D**) Corresponding bottom and side views of the H-ficolin structure. The side chains of the binding site residues are displayed as ball and sticks and colored green (S1), red (S2), black (S3), and orange (S4). To enhance clarity of the side view, only one of each representative binding sites is shown on the L-ficolin trimer. N and C indicate the N- and C-terminal ends of each protomer. Ca^{2+} ions are represented as golden spheres. Figure generated using MOLSCRIPT (Kraulis, 1991).

statistics are listed in Supplementary Table II. As expected from the trimeric nature of ficolins (Matsushita and Fujita, 2001) and in agreement with our ultracentrifugation analysis of L-ficolin, both recognition domains formed homotrimers in all crystal forms. The protomers associate as a three-lobed clover-like structure with pseudo-three-fold symmetry (Figure 1). The assemblies have a small central cavity and are rather open, with clefts separating the upper parts of the protomers. Their overall dimensions are 50–55 Å in height and 80 Å in maximum diameter. The N-termini of the three protomers emerge at the base of the trimers (Figure 1B), consistent with the fact that they are connected to a collagen-like triple helix in the intact molecules.

In both proteins, a Ca^{2+}-binding site is found in each protomer, at a position homologous to that of TL5A (Kairies *et al*, 2001), that is, in a loop region that represents the most external part of the trimers. The distances between the Ca^{2+} ions are approximately 65 Å for L-ficolin and 67 Å for H-ficolin (Figure 1A and C). In agreement with mass spectrometry analyses, of the two potential N-glycosylation sites of L-ficolin (Asn215 and Asn275), only the former is occupied by an oligosaccharide chain. In the most favorable case, a segment comprising two *N*-acetyl-glucosamines with a fucose bound to the proximal one and two additional mannoses could be built into well-defined electron density. The single N-glycosylation site of H-ficolin at Asn166 is clearly also linked to an oligosaccharide, but this could not be defined owing to ill-defined electron density. In both proteins, the carbohydrates are located at the periphery of the trimers.

Protomer topology

The L- and H-ficolin protomers have an ellipsoidal structure similar to that of TL5A and are subdivided into the same three domains (Kairies *et al*, 2001) (Figure 2). In the N-terminal A domain (residues Pro70–Gly116 in L-ficolin), a major difference with TL5A is that the loop connecting strands β1 and β2 is shorter and does not exhibit an α-helical conformation (Figure 2A–C). Additionally, this loop is anchored to the N-terminal segment of the domain through a disulfide bond (Cys73–Cys101 in L-ficolin, Cys63–Cys87 in H-ficolin) that is not present in TL5A. The other two disulfide bonds seen in the ficolin structures (Cys80–Cys108 and Cys232–Cys245 in L-ficolin) are homologous to those observed in TL5A (Figure 2).

The larger domain B (residues Trp117–Ala206 and Lys279–Ala288 in L-ficolin) is clamped between domains A and P (Figure 2A and B) and comprises the seven-stranded anti-parallel β-sheet and the two helices (α4 and α5) already

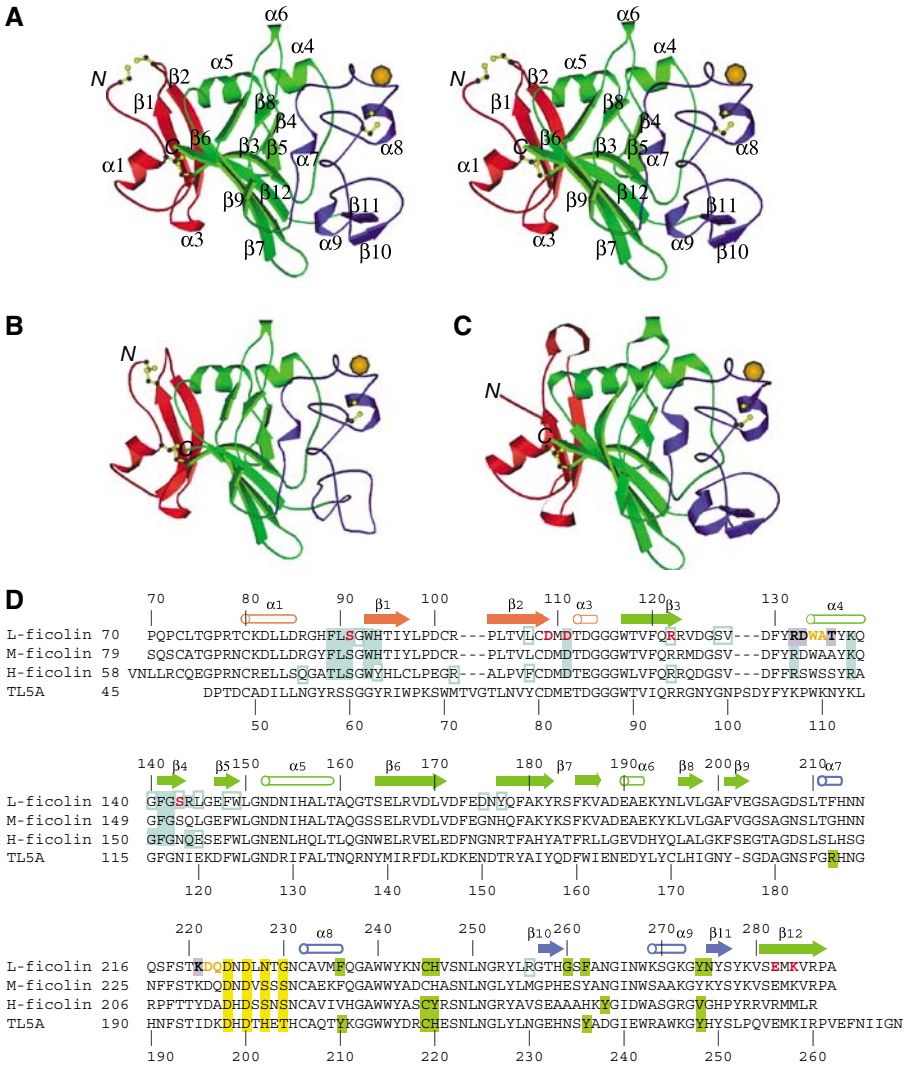

Figure 2 Structure of the fibrinogen-like protomers. (**A**) Stereo view of the L-ficolin protomer. (**B, C**) Comparative views of the H-ficolin and TL5A protomers in the same orientation. Domains A, B, and P are colored red, green, and blue, respectively. Disulfide bonds are shown in yellow. N and C indicate the N- and C-terminal ends. Figure generated using MOLSCRIPT. (**D**) Sequence alignment of the fibrinogen-like domains of L-ficolin, M-ficolin, H-ficolin, and TL5A. Numbers on top and below refer to the L-ficolin and TL5A sequences, respectively. Secondary structure elements are shown in red (A domain), green (B domain), and blue (P domain). The binding site residues are colored using the same code as in Figure 1: green (S1), bold red (S2), black (S3), and bold orange (S4). Residues involved in the inter-subunit interfaces are colored cyan, variable ones being boxed. Residues involved in Ca^{2+} binding in L-ficolin, H- ficolin, and TL5A are displayed in a yellow box.

described in TL5A (Kairies *et al*, 2001). Comparison of L- and H-ficolins with TL5A reveals, however, several significant differences, notably in the loop connecting strand β3 to helix α4 (Figure 2A–C). The C-terminal segment following strand β12 is also much shorter in the ficolins than in TL5A (Figure 2D).

Domain P of L-ficolin (residues Gly207–Tyr278) comprises three short α-helices (α7–α9) and a short antiparallel two-stranded β-sheet (strands β10, β11) similar to those observed in TL5A. In H-ficolin, in contrast, the region corresponding to β10, α9, and β11 has a strikingly different sequence and shows no defined secondary structure. In this region, the maximal displacement between the polypeptide chains of L- and H-ficolins (about 8 Å) occurs at residues Ser250 (H-ficolin) and Ser260 (L-ficolin). In both proteins, a Ca^{2+} ion is bound to the large loop connecting helices α7 and α8. Coordination in L-ficolin involves two water molecules, both carboxylate oxygens of Asp224, one of the side-chain oxygens of Asp226, and the main-chain carbonyl oxygens of Asn228 and Gly230. Similarly, the Ca^{2+} coordination in H-ficolin is mediated by two water molecules and residues Asp214, Asp216, Ser218, and Ser220, which contribute ligands in the same way as in L-ficolin. The average bond distance is 2.4 Å, a characteristic value for known Ca^{2+}-binding sites in proteins (Harding, 1999). These sites are very similar to the one described in TL5A, which involves homologous residues (Asp198, Asp200, His202, and Thr204, respectively) (Figure 2D; Kairies *et al*, 2001).

Compared to TL5A, domain A of the ficolins is slightly rotated with respect to the remainder of the structures, with a rotation of approximately 6° about the loop connecting helix α3 to strand β3 (Figure 2A). The root mean square deviation (r.m.s.d.) between the two ficolins is 0.74 Å, based on 201 Cα atoms, and of about 0.6 Å when domains A, B, and P are superimposed individually. For comparison, the r.m.s.d.

between L-ficolin and the fibrinogen γ fragment (Yee *et al*, 1997) is 1.0 Å, based on 204 Cα atoms.

Analysis of inter-subunit interfaces reveals structural plasticity

In both ficolins, the inter-protomer interfaces mainly involve homologous residues located before or within strand β1 on one side and residues contributed by helix α4 and strands β4 on the other. These residues provide a framework common to all interfaces in both proteins, consisting of two hydrogen bonds between the main chain of Ser91 and the main chain of Gly142 and Arg144, a salt bridge between Asp111 and Arg132, and van der Waals contacts between Phe89 and Gln139, and between Leu90 and the main chain of Phe141 and/or Gly142 (L-ficolin numbering). There are several other hydrophobic contacts, such as the one between the side chains of Leu90 and Trp93 of L-ficolin (Leu80 and Trp83 in H-ficolin), which is variable in the former case, but conserved in the latter. H-ficolin also features one additional hydrogen bond between the side chain of Thr79 and the main chain of Ala129. In both proteins, the total buried surface at the three interfaces is approximately 2900 $Å^2$. Most of the interface residues of L- and H-ficolins are conserved in M-ficolin, whereas more significant differences are found in TL5A, notably in the regions corresponding to segments Phe89–His94 and Arg144–Gly146 of L-ficolin (Figure 2D).

A striking characteristic of the L-ficolin trimer is the variability of the angle between the A, B, and C protomers. Of the various structures investigated, a single one (L-9) has a true three-fold symmetry axis, with angles of 120° between the protomers. In all other cases, the interfaces, particularly A/B and A/C, are stabilized by a variable number of van der Waals contacts. As a result, the inter-protomer angles vary from 117 to 122°, with maximum rotation amplitudes for protomer A of L-1 and L-6, which diverge from each other by 14°. Interestingly, these rotations do not disrupt the common interaction framework described above, underlying the remarkable plasticity of the inter-protomer interfaces.

Location of the major ligand-binding sites observed in L- and H-ficolins

Several ligands were allowed to bind to the recognition domains of L- and H-ficolins in order to locate their binding site(s) and investigate their mode of interaction. This was achieved by co-crystallization and/or by soaking crystals into ligand-containing solutions. The refined structure obtained in each case was then compared with the native one, and ligand location was determined from additional electron density.

In the case of H-ficolin, binding was observed only when using galactose and D-fucose, and both carbohydrates were found to bind to site S1. This site lies on the outer side of the trimer, in the vicinity of the Ca^{2+}-binding site, and is homologous to the GlcNAc-binding site identified in TL5A (Figures 1 and 2). This external site involves hydrophobic and aromatic residues, namely Cys235, Tyr236, Tyr254, and Val264.

A more extensive structural study using various carbohydrate and non-carbohydrate ligands was performed on L-ficolin (Table I), allowing us to identify three additional binding sites, S2–S4, which together define a new recognition surface, extending on both sides of the cleft in between the upper parts of the protomers (Figure 1). Most of the ligands tested bound at the major sites S2 and/or S3, mainly involving polar and charged residues. The inner site S2 is located close to the inter-protomer interface, and comprises residues Ser91, Asp109, Asp111, Arg122, Glu282, and Lys284 from one protomer and Ser143 from the adjacent protomer. S3 lies on the opposite side of the cleft, at a position intermediate between S2 and the Ca^{2+}-binding site, and involves residues Arg132, Asp133, Thr136, and Lys221. S4 is located between S3 and the Ca^{2+}-binding site (Figure 1) and comprises both hydrophobic (Trp135, Ala136) and polar (Asp222, Gln223) residues. In our study, it appears to be a minor binding site, as only two different carbohydrate residues were observed at this location.

The external S1 binding site in TL5A and the ficolins

S1 is homologous to both the polymerization pocket of the fibrinogen γ fragment and the GlcNAc-binding pocket of TL5A (Kairies *et al*, 2001). Consequently, it was expected to be involved in ligand recognition in the ficolins. However, comparative analysis of the structures of S1 in TL5A and ficolins L and H reveals some common features but also significant variations in the nature and orientation of the side chains (Figures 2 and 3); thus, these sites differ in ligand specificity. In TL5A, a funnel is delineated by the aromatic residues Tyr210, Tyr236, Tyr248, and His220, the methyl side chain of Ala237 being located at the base (Kairies *et al*, 2001; Figure 3B). This funnel is essentially conserved in L-ficolin

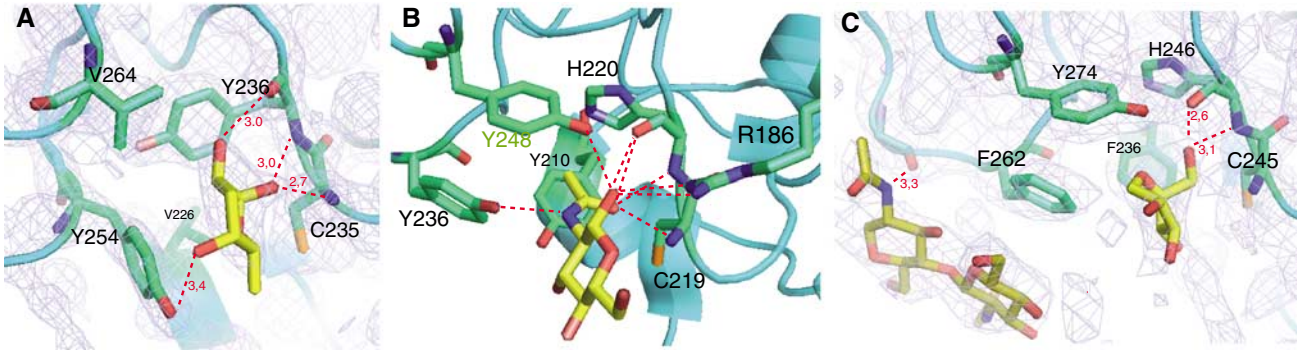

Figure 3 Comparative views of the S1 binding site in H-ficolin, L-ficolin, and TL5A. The side chains of the residues defining S1 are colored green and the ligands are displayed in yellow. (**A**) D-Fucose bound to H-ficolin. (**B**) GlcNAc bound to TL5A. (**C**) The terminal mannose of the oligosaccharide chain from a neighboring molecule positioned in site S1 of L-ficolin. On the left, the two proximal GlcNAc residues of the chain interacting on the edge of the binding site.

(Figure 3C), but it is deeper and reorganized on one side in H-ficolin (Figure 3A). The unusual *cis* conformation of Cys219 seen in TL5A is also found in the homologous residues Cys245 and Cys235 of ficolins L and H. Thus, in each case, the backbone NH group of this Cys residue and the backbone atoms of the following amino acid are properly oriented to provide tight hydrogen bonds to a ligand introduced into the funnel. Nevertheless, whereas in TL5A S1 appears specifically designed to bind GlcNAc, this is not true in the ficolins. Two major structural differences account for this: (i) Tyr236, which in TL5A stabilizes the acetamido nitrogen of GlcNAc through its hydroxyl group, is not conserved in the ficolins; (ii) similarly, the side chains of Arg186 and Tyr210, which participate in the recognition of the sugar ring in TL5A, are not present in the ficolins (Figure 3).

In the case of H-ficolin, the best defined electron density in site S1 was observed using D-fucose as a ligand (Figure 3A). The interaction mainly involves three H-bonds between the conserved backbone atoms of the Cys235–Tyr236 segment mentioned above, and two consecutive hydroxyl groups (1-OH and 2-OH) of the sugar ring. Another interaction, involving also the Cys235–Tyr236 segment, was observed for galactose. In contrast, no interaction could be detected using L-fucose, in full agreement with previous competition experiments (Sugimoto *et al*, 1998).

In the case of L-ficolin, none of the tested ligands was observed in S1. In contrast, the edge of site S1 was consistently involved in the packing of crystal form A by binding the two proximal GlcNAc residues of the oligosaccharide chain of a neighboring molecule (Figure 3C), through hydrogen bonds and van der Waals contacts with residues Gly260 and Phe262. Very well-defined electron density was always observed for these interacting carbohydrate residues. In some cases, an additional electron density, interpretable as a terminal mannose residue from this chain, was observed inside the S1 funnel (Figure 3C). Again, the interaction involved two H-bonds between the backbone atoms of His246 and the 6-OH group of the mannose ring. These packing interactions with carbohydrates in form A of L-ficolin likely reflect a significant affinity for oligosaccharides in this area. Further investigations will be required to establish its physiological significance.

L-ficolin recognizes galactose and 1,3-β-D-glucan in a highly specific manner

Very well-defined electron densities were observed in the other sites for two neutral carbohydrates ligands, galactose and the linear oligosaccharide 1,3-β-D-glucan, likely reflecting a single, highly specific recognition mode for each ligand. Galactose was tested as a ligand for L-ficolin after analyses by surface plasmon resonance spectroscopy indicated that both recombinant human L-ficolin (Hummelshoj *et al*, 2007) and purified human L-ficolin (NM Thielens, unpublished data) bind to immobilized galactose–BSA with high affinity ($K_D = 0.8$ nM). Galactose mostly binds to S2, although minor binding to S4 as part of crystal packing interactions was also observed (Table I). All the oxygen atoms of the carbohydrate ring are stabilized in site S2 through one or two hydrogen bonds (Figure 4A). These bonds involve the carboxyl groups of Glu282 and Asp109 (2-OH, 3-OH, and 4-OH), the hydroxyl groups of Ser91 (4-OH) and Ser143 (O5), and the carboxyl group of Asp111 (6-OH). The 1-OH group interacts with a water molecule that is oriented toward the solvent channel, in a position that mimics a link to a carbohydrate chain (Figure 4A). Consequently, it is tempting to postulate that site 2 provides a recognition pocket for the terminal galactose residue of a linear oligosaccharide chain.

Another nicely defined electron density was obtained for the linear oligosaccharide β-D-glucan (Figure 4G and H), composed of four glucose units linked through covalent β-1-3 bonds. L-ficolin was shown to be the only serum protein able to bind specifically to 1,3-β-D-glucan, a molecular marker of yeast and fungal cell walls (Ma *et al*, 2004). It should be emphasized that this is the first experimental description of the interface between a 1-3-β-D-glucan and a recognition protein. The first three glucose units of the molecule are roughly arranged in a linear fashion, and bind to L-ficolin through interactions spanning sites S3 and S4, forming a network of direct and water-mediated hydrogen bonds tightly stabilizing each of their 2-OH groups (Figure 4G). A curvature in the molecule then allows Glc4 to bind to site S4 through a stacking interaction with the side chain of Trp134 (Figure 4H). Further stabilization of the 4-OH group of Glc4 is provided by a network of hydrogen bonds involving the side-chain carbonyl group of Gln223 and a Ca^{2+} ion, mainly coordinated by water molecules, bound at the edge of site S4 (Figure 4H).

L-ficolin binds to a wide range of small acetylated ligands in a variety of orientations

Various acetylated compounds were also tested as ligands of L-ficolin (Table I). GlcNAc binding was initially described for L-ficolin (Matsushita *et al*, 1996, 2000; Le *et al*, 1998), and K_D values in the low nanomolar range were determined for both recombinant and purified human L-ficolin using GlcNAc–BSA (Hummelshoj *et al*, 2007; NM Thielens, unpublished data). In addition, a variety of N-acetylated carbohydrates (GlcNAc, GalNAc, ManNAc) and other non-carbohydrate acetylated compounds (GlyNAc, CysNAc, acetylcholine) were shown to inhibit efficiently the binding of L-ficolin to GlcNAc–Sepharose beads and to *Streptococcus pneumoniae* (Krarup *et al*, 2004).

Binding of only two acetylated compounds, *N*-acetylcysteine (CysNAc) and GlcNAc, was observed in site S2 (Figure 4B and C) and the acetyl group was not recognized the same way in the two cases. Moreover, different orientations could be modelled for the same ligand on different protomers or in different crystal forms. Indeed, the particular topology of site S2, with charged or polar side chains arranged in a ring-like fashion (Figure 4B and C), provides alternative interaction sites for the N-acetyl group, resulting in the observed binding versatility of these small ligands. Moreover, the shape of the S2 cavity can be very slightly modulated by the above-mentioned plasticity at the inter-protomer interfaces.

Most of the acetylated ligands tested bind to site S3 (Figure 4D–F). As the acetyl group is the only common feature of these molecules, its binding would be expected to be similar in all cases. However, although the acetyl moiety of all these ligands binds to the backbone nitrogen of Asp133, indicating some sort of specificity, the interaction is not strictly conserved (Figure 4F). Thus, the different molecules are observed in a variety of orientations, these small ligands lacking any other common mode of interaction. Acetylcholine is an exception to this pattern because a well-defined electron

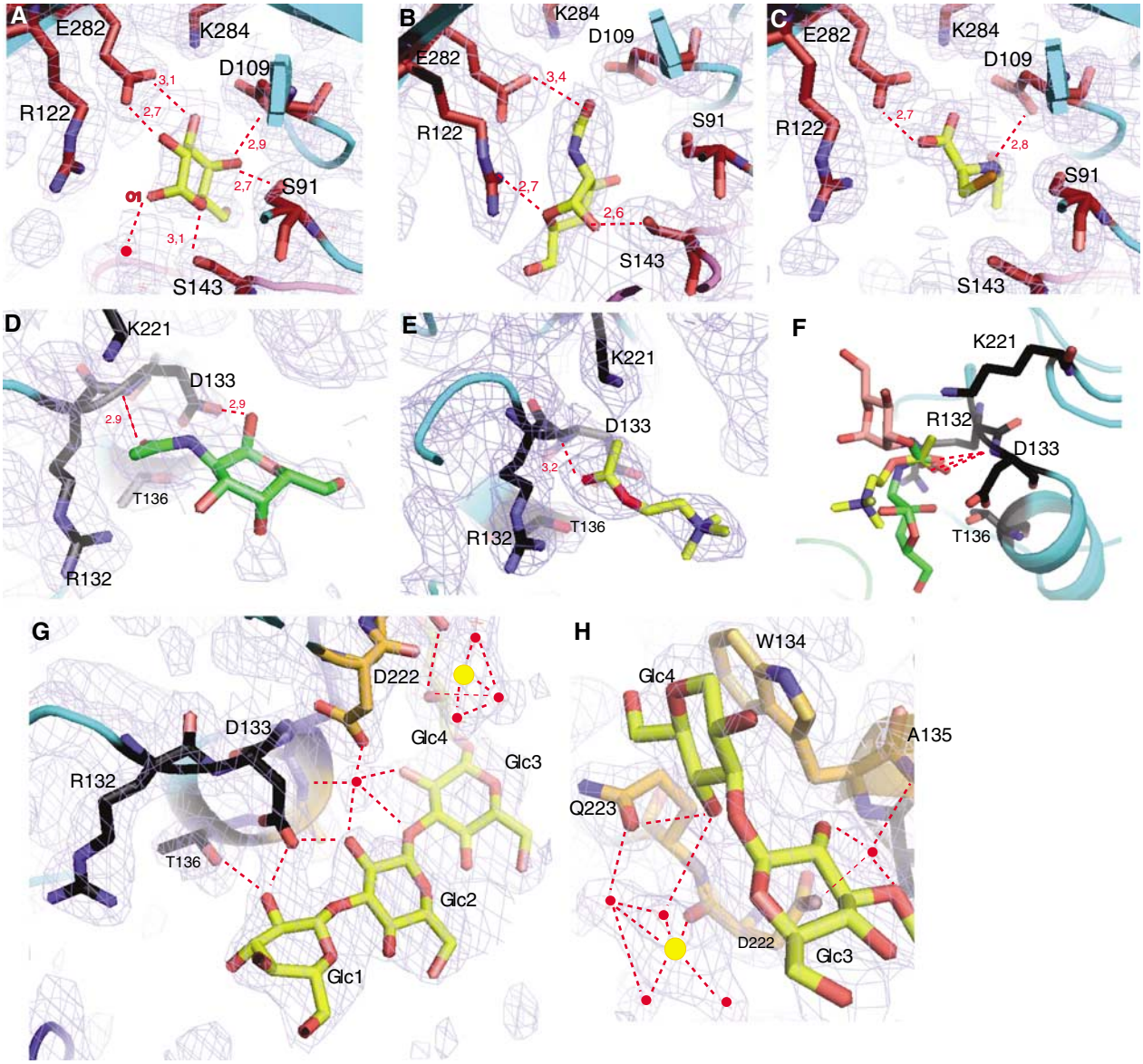

Figure 4 The novel binding sites identified in L-ficolin. (**A–C**) Interaction of galactose, GlcNAc, and N-acetylcysteine (CysNAc) in site S2. The interaction of galactose O6 with D111, at the back of the binding pocket, is not visible on this orientation. (**D**) Detailed view of the interaction of GalNAc in S3. (**E**) Binding of acetylcholine to site S3. (**F**) Superimposition of different acetylated ligands bound to site S3. Three molecules have been selected for clarity purposes, showing both medium and extreme orientations. The common acetyl group interaction is highlighted with dashed lines. Acetylcholine, GalNAc, and ManNAc are shown in yellow, green, and pink, respectively. A more extensive superimposition of the different acetylated molecules bound to site S3 is provided as Supplementary data. (**G**) Interaction of the first three glucose units of 1,3-β-D-glucan with residues from sites S3 (black) and S4 (orange). (**H**) A different view of the binding of 1,3-β-D-glucan, showing interaction of Glc3 and Glc4 with residues from sites S3 (black) and S4 (orange). Polar interactions are represented by dashed red lines. Water molecules are shown as red dots and the calcium ion is represented by a yellow sphere.

density, defining a single orientation, was observed at low ligand concentrations (Figure 4E). This is most likely due to the relative rigidity of this molecule compared to acetylated carbohydrates such as GalNAc (Figure 4D).

Thus, in sharp contrast with the highly specific GlcNAc recognition observed in TL5A (Kairies *et al*, 2001), the established specificity of L-ficolin for acetylated compounds is not explained by a single mode of interaction at the atomic level. Instead, recognition of these molecules appears to involve minimal structural requirements for the acetyl group, and virtually no requirement for the remainder of the molecule. Nevertheless, this seems consistent with the ability of L-ficolin to accommodate a wide range of small acetylated compounds (Krarup *et al*, 2004). This is also likely to reflect the ability of L-ficolin to accommodate, through different subsites, acetylated markers (units) within larger physiological ligands. In this regard, it is interesting to note that the first glucose residue of the elongated β-glucan molecule binds to the edge of site S3, a position that could also accommodate a GlcNAc residue with its acetyl group stabilized into this site (Figure 4G and D).

Ficolins as novel types of pattern recognition molecules

After C1q and MBL, L- and H-ficolins provide additional examples of multivalent recognition molecules that sense molecular patterns on microbial or apoptotic surfaces by

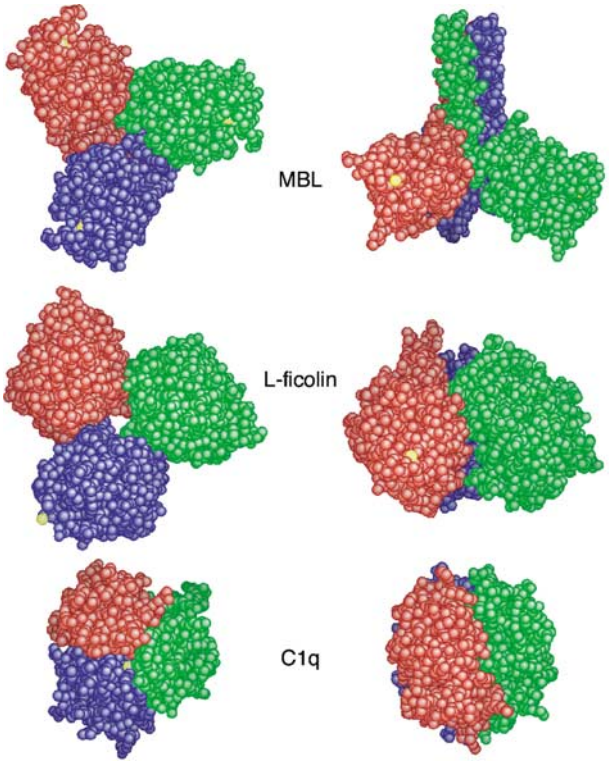

Figure 5 Comparative views of the trimeric recognition domains of MBL, L-ficolin, and C1q. Space-filling bottom views (left) and side views (right) are shown. Ca^{2+} ions are represented as yellow spheres. Figure generated using GRASP (Nicholls *et al*, 1991).

means of a trimeric ligand-binding domain. As illustrated in Figure 5, the semi-open ficolin structure is intermediate between the compact assembly seen in the complement protein C1q (Gaboriaud *et al*, 2003), where subunits associate through an extensive interface (total buried surface = 5490 Å^2), and the widely open structure of the MBL carbohydrate recognition domain (Sheriff *et al*, 1994; Weis and Drickamer, 1994), which trimerizes through a triple α-helical neck region, with minimal interaction between the carbohydrate recognition domains (buried surface = 426 Å^2). In this respect, it is noteworthy that, contrary to what was generally believed based on amino-acid sequence analyses (Fujita, 2002), L-and H-ficolins do not contain a neck region. This likely also applies to M-ficolin, which, as judged from conservation of the residues involved in inter-subunit interfaces (Figure 2), is expected to retain the trimeric structure observed in L- and H-ficolins.

Whereas the recognition properties of MBL involve highly specific hydrogen bonding interactions with only the hydroxyl groups at positions C3 and C4 of a terminal carbohydrate residue (Weis *et al*, 1992), a pronounced plasticity of recognition is observed in the case of L-ficolin. By investigating the structural basis for this unusual recognition versatility, we have identified a novel extended recognition area, located around the cleft between the ficolin subunits. This area can be subdivided into different contiguous subsites (S2–S3–S4). The resulting binding groove is reminiscent of the one observed in the peptidoglycan recognition proteins (Chang *et al*, 2006; Lim *et al*, 2006), and thus seems appropriately shaped to bind extended polysaccharides. This is illustrated by the structure of the complex between L-ficolin and 1,3-β-D-glucan, which reveals how the tetrasaccharide is coordinated by a series of interactions involving residues contributed by both S3 and S4.

Interestingly, each of these recognition proteins seems to complement the recognition spectrum of the others. For example, the external S1 site in H-ficolin will bind galactose or D-fucose, which are not recognized by MBL. On the other hand, GlcNAc is a marker that can be recognized by both MBL and L-ficolin, but in different contexts, namely at the tip or inside elongated oligosaccharide molecules, respectively.

Implications for the recognition of microbial and apoptotic surfaces by ficolins

Although the precise molecular nature of the physiological ligands of ficolins remains poorly characterized, especially in the case of H-ficolin, the present study brings new insights into the molecular markers they can recognize. Obviously, if these markers are exposed onto the target surfaces, they will contribute to the recognition process. Conversely, their absence will prevent or reduce binding. In this respect, it is noteworthy that D-fucose was recently identified in a bacterial surface layer glycan (Kählig *et al*, 2005). Thus, the very narrow spectrum of microbial targets identified for H-ficolin (Krarup *et al*, 2005) may be related to its restricted specificity toward galactose and D-fucose, and not to L-fucose or to acetylated carbohydrates, which are frequent microbial markers.

L-ficolin has been clearly implicated in the opsonophagocytosis of type III group B streptococci (Aoyagi *et al*, 2005). The polysaccharide capsular antigen of the latter shows an elongated structure, comprising GlcNAc residues not in a terminal position, but also various β-1-3 and β-1-4 links between glucose and galactose rings (Kadirvelraj *et al*, 2006). The present study strongly suggests that L-ficolin is designed to bind such elongated carbohydrates, presenting a number of acetylated and neutral residues. Another example arises from the known specific binding of L-ficolin to lipoteichoic acid, a cell wall component found in all Gram-positive bacteria (Lynch *et al*, 2004), which is also an elongated GlcNAc-containing molecule. Like other 1,3-β-D-glucan-recognizing proteins, L-ficolin is likely to bind to the surfaces of yeast or other fungal cells, but also to some bacterial surfaces (Wang *et al*, 2006). Moreover, as the interaction between L-ficolin and 1,3-β-D-glucan does not involve the C4 hydroxyl groups of the latter, a similar interaction could also be extrapolated to poly-1,3-β-D-galactosyl polysaccharide determinants, which are for instance virulence factors on *Leishmania major* parasites (Pelletier *et al*, 2003). In the same way, it would be interesting to investigate whether L-ficolin is able to bind to plasmodium surface glycoproteins, which expose a terminal galactose (Ramasamy and Reese, 1986).

As for the recognition of apoptotic cells by ficolins (Kuraya *et al*, 2005), it is interesting to note that recent studies have described the specific exposure at the apoptotic cell surface of phagocytic markers such as galactose (Bilyy *et al*, 2004; Yuita *et al*, 2005) or GlcNAc (Orlando and Pittman, 2006), which are shown here to be ligands of H- or L-ficolins.

In conclusion, the present study reveals how L- and H-ficolins have evolved different binding sites, explaining their very different recognition spectra. We have discovered an

extended binding surface area on L-ficolin and shown the recognition plasticity of some of its subsites. Based on the nature of its known physiological ligands, we conclude that L-ficolin functions as a sensor of acetylated and neutral carbohydrate markers within extended polysaccharide structures.

Materials and methods

Protein expression and purification

The DNA segments encoding residues 70–288 and 58–276 of mature human L- and H-ficolins were amplified using $Vent_R$ polymerase and the PCRBac-P35 and pGEM-T easy plasmids containing the full-length cDNAs (Matsushita *et al*, 1996) as templates, according to established procedures. Both segments start at the first residue following the collagen-like sequence and end at the C-terminus of the protein. The amplified DNAs were cloned in-frame with the melittin signal peptide of the pNT-Bac baculovirus transfer vector (Rossi *et al*, 1998), and the recombinant baculoviruses were generated using the Bac-to-Bac^{TM} system (Invitrogen Corp.), amplified, and titrated as described previously (Thielens *et al*, 2001). High Five cells were infected with the recombinant viruses for 96 h at 27°C. L-ficolin was purified from the culture supernatant by ion-exchange chromatography on a Q-Sepharose Fast Flow column (Amersham Biosciences) equilibrated in 50 mM triethanolamine–HCl, pH 8.0, using a linear gradient to 250 mM NaCl. H-ficolin was passed through a DEAE-cellulose DE52 column (Whatman) equilibrated in 5 mM $CaCl_2$ and 50 mM triethanolamine–HCl, pH 8.0. The flow-through fraction was dialyzed against 5 mM $CaCl_2$ and 50 mM MES, pH 5.8, and further purified by ion-exchange chromatography on an S-Sepharose Fast Flow column (Amersham Biosciences) equilibrated in the same buffer, using a linear NaCl gradient as described above.

Ultracentrifugation and mass spectrometry analyses

Sedimentation velocity analysis was performed at 42 000 r.p.m. and 20°C using a Beckman XL-I analytical ultracentrifuge as described previously (Zundel *et al*, 2004). The sedimentation (*s*) and diffusion (*D*) coefficients were estimated using the SEDFIT program (http://www.analyticalcentrifugation.com) as described by Lacroix *et al* (2001). The partial specific volume was estimated from the amino-acid composition at 0.714 ml/g using the SEDNTERP program (http://www.rasmb.bbri.org), which was also used for calculation of the corrected $s_{20,w}$ value. Mass spectrometry analysis was performed using the matrix-assisted laser desorption ionization technique on a voyager Elite XL instrument (PerSeptive Biosystems, Cambridge, MA) under conditions described previously (Lacroix *et al*, 1997).

Crystallization and data collection

Recombinant L-ficolin was concentrated to 5.7 mg/ml in 145 mM NaCl and 50 mM triethanolamine–HCl, pH 7.4. Native crystals (form A) suitable for X-ray data collection were obtained at 20°C by the hanging drop vapor diffusion method by mixing equal volumes (2 µl) of the protein solution and of a reservoir solution composed of 15% (w/v) PEG 8000, 200 mM Ca acetate, and 0.1 M Hepes, pH 7.0. Cryoprotection was achieved by adding 1 µl of PEG 400 to the drop, just before flash-cooling the crystal in liquid nitrogen.

Most of the L-ficolin–ligand complexes listed in Table I were obtained in the A crystal form, by either co-crystallization and/or soaking experiments. The best diffracting crystals (L-2, 1.95 Å resolution) were obtained by co-crystallization with CysNAc. Microseeding experiments were then performed using these crystals to generate native crystals of higher and more reproducible quality. These native crystals were then soaked overnight in 150 mM ligand-containing solutions to obtain the galactose (L-3), β-glucan (L-4), and CysNAc (L-10) complexes X-ray data sets. L-ficolin–GlcNAc complexes were obtained in two other crystal forms. These were prepared by adding 10 mM GlcNAc and 1 mM $CaCl_2$ to the concentrated protein solution. The reservoir solution contained 0.9–0.95 M sodium succinate, pH 6.5 (form C) or pH 7.0 (form B), and 1 mM phosphoglycerate in the case of form C. Cryoprotection solutions were 1.05 M Na succinate, pH 7.0, 10 mM GlcNAc, 10 mM Ca acetate, 24% glycerol (form B) and 1 M Na succinate, pH 6.5, 1 mM phosphoglycerate, 22% glycerol (form C). Several L-ficolin crystals were obtained in a fourth crystal form D, of space group $P222_1$ ($a = 109$ Å, $b = 130$ Å, $c = 268$ Å, $\alpha = \beta = \gamma = 90°$) with four trimers per asymmetric unit.

H-ficolin was concentrated to 3 mg/ml in the presence of 100 mM non-detergent sulfobetain 195. The hanging drop vapor diffusion method was used with a reservoir containing 21–26% PEG 6000, 0.1 M Tris–HCl pH 8 or 8.5, and 20 mM $CaCl_2$. The ligands were introduced also in the reservoir solution in increasing concentrations up to 25 mM for galactose and 50 mM for fucose. Cryoprotection was achieved using paratone oil (crystal form H-4) or by adding 20% glycerol (H-1 to H-3).

Measurements were performed at ESRF beamlines ID14, ID23, ID29, BM16, and BM30 (see Supplementary Table II). Diffraction images were processed using the program XDS (Kabsch, 1993). Complete crystallographic data statistics are listed in Supplementary Table II. The different cell parameters and space groups are also provided in Table I.

Structure determination and refinement

The structures of L-ficolin and H-ficolin were solved by molecular replacement with AMoRe (Navaza, 2001), first using the TL5A structure (PDB code 1JC9) as a search model to solve the L-ficolin structure and then using the latter. Model rebuilding was performed using the graphics programs O (Jones *et al*, 1991) and Coot (Emsley and Cowtan, 2004). Initial refinement of the first L-ficolin models was carried out with CNS (Brünger *et al*, 1998). The last refinement steps were carried out with Refmac5 (Murshudov *et al*, 1997), using the Arp procedure (Lamzin and Wilson, 1997) to introduce the water molecules. The quality of the map allowed construction of at least 213 residues (segment 76–288) out of the 219 residues of the recombinant fragment (70–288). The N-terminal extremity of the H-ficolin fragment 58–276 is disordered, the model starting at residues 63–66. The N-terminal extremity exhibits different conformations depending on the crystal environment, and was only fully defined in molecule A of the L-ficolin B crystal form. Residue Cys245 and Cys235 are in the *cis* configuration in L- and H-ficolin, respectively. TLS refinement was used for most of the L-ficolin models, and non-crystallographic symmetry constraints were introduced for data sets at a resolution lower than 2.5 Å. Molecule A in crystal form A was often partially ill-defined in its most external part because of its subtle rotational freedom in the trimer (see text about the structural plasticity of the trimeric interface). The diffraction data sets obtained with the D crystal form were used to check qualitatively the position and content of the binding sites in a different crystal packing environment but the corresponding models were not extensively refined. Consequently, they have not been included in Table I.

Data deposition

The atomic coordinates for crystal nos. H-1, H-2, H-3, and L-1 to L-10 have been deposited in the Protein Data Bank (www.rcsb.org), as listed in Table I.

Supplementary data

Supplementary information is available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Jordi Juanhuix for his contribution to the initial stages of this work, Christine Ebel and Bernard Dublet for their help with ultracentrifugation and mass spectrometry analyses, Jean-Baptiste Reiser for scientific and computing advices, Stéphanie Combe for generating recombinant baculoviruses, and the beamline scientists at the ESRF for their help with the use of X-ray diffraction equipment. This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, the Université Joseph Fourier, Grenoble, and the Agence Nationale de la Recherche (grant ANR-05-MIIM-023-01).

References

Adachi Y, Ishii T, Ikeda Y, Hoshino A, Tamura H, Aketagawa J, Tanaka S, Ohno N (2004) Characterization of beta-glucan recognition site on C-type lectin, dectin 1. *Infect Immun* **72:** 4159–4171

Aoyagi Y, Adderson EE, Min JG, Matsushita M, Fujita T, Takahashi S, Okuwaki Y, Bohnsack JF (2005) Role of L-ficolin/mannose-binding lectin associated serine protease complexes in the opsonophagocytosis of type III group B streptococci. *J Immunol* **174:** 418–425

Bilyy RO, Antonyuk VO, Stoika RS (2004) Cytochemical study of role of α-D-mannose- and β-D-galactose-containing glycoproteins in apoptosis. *J Mol Histol* **35:** 829–838

Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54:** 905–921

Chang CI, Chelliah Y, Borek D, Mengin-Lecreulx D, Deisenhofer J (2006) Structure of tracheal cytotoxin in complex with a heterodimeric pattern-recognition receptor. *Science* **311:** 1761–1764

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60:** 2126–2132

Frederiksen PD, Thiel S, Larsen CB, Jensenius JC (2005) M-ficolin, an innate immune defence molecule, binds patterns of acetyl groups and activates complement. *Scand J Immunol* **62:** 462–473

Fujita T (2002) Evolution of the lectin-complement pathway and its role in innate immunity. *Nat Rev Immunol* **2:** 346–353

Gaboriaud C, Juanhuix J, Gruez A, Lacroix M, Darnault C, Pignol D, Verger D, Fontecilla-Camps JC, Arlaud GJ (2003) The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *J Biol Chem* **278:** 46974–46982

Guan R, Roychowdury A, Ember B, Kumar S, Boons GJ, Mariuzza RA (2005) Crystal structure of a peptidoglycan recognition protein (PGRP) in complex with a muramyl tripeptide from Gram-positive bacteria. *J Endotoxin Res* **11:** 41–46

Harding M (1999) The geometry of metal–ligand interactions relevant to proteins. *Acta Crystallogr D* **55:** 1432–1443

Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RAB (1999) Phylogenetic perspectives in innate immunity. *Science* **284:** 1313–1318

Holmskov U, Thiel S, Jensenius JC (2003) Collectins and ficolins: humoral lectins of the innate immune defense. *Annu Rev Immunol* **21:** 547–578

Hummelshoj T, Thielens NM, Madsen HO, Arlaud GJ, Sim RB, Garred P (2007) Molecular organization of human ficolin-2. *Mol Immunol* **44:** 401–411

Janeway CA (1992) The immune system evolved to discriminate infectious nonself from non-infectious self. *Immunol Today* **13:** 11–16

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47:** 110–119

Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* **26:** 795–800

Kadirvelraj R, Gonzalez-Outeirino J, Foley BL, Beckham ML, Jennings HJ, Foote S, Ford MG, Woods RJ (2006) Understanding the bacterial polysaccharide antigenicity of *Streptococcus agalactiae* versus *Streptococcus pneumoniae*. *Proc Natl Acad Sci USA* **103:** 8149–8154

Kählig H, Kolarich D, Zayni S, Scheberl A, Kosma P, Schaffer C, Messner P (2005) N-acetylmuramic acid as capping element of alpha-D-fucose-containing S-layer glycoprotein glycans from *Geobacillus tepidamans* GS5-97 T. *J Biol Chem* **280:** 20292–20299

Kairies N, Beisel HG, Fuentes-Prior P, Tsuda R, Muta T, Iwanaga S, Bode W, Huber R, Kawabata SI (2001) The 2.0-Å crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems. *Proc Natl Acad Sci USA* **98:** 13519–13524

Kim J-I, Lee CJ, Jin MS, Lee C-H, Paik S-G, Lee H, Lee J-O (2005) Crystal structure of CD14 and its implications for lipopolysaccharide signalling. *J Biol Chem* **280:** 11347–11351

Krarup A, Sorensen UB, Matsushita M, Jensenius JC, Thiel S (2005) Effect of capsulation of opportunistic pathogenic bacteria on binding of the pattern recognition molecules mannan-binding lectin, L-Ficolin and H-Ficolin. *Infect Immun* **73:** 1052–1060

Krarup A, Thiel S, Hansen A, Fujita T, Jensenius JC (2004) L-ficolin is a pattern recognition molecule specific for acetyl groups. *J Biol Chem* **279:** 47513–47519

Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* **24:** 946–950

Kuraya M, Ming Z, Liu X, Matsushita M, Fujita T (2005) Specific binding of L-ficolin and H-ficolin to apoptotic cells leads to complement activation. *Immunobiology* **209:** 689–697

Lacroix M, Ebel C, Kardos J, Dobo J, Gal P, Zavodszky P, Arlaud GJ, Thielens NM (2001) Assembly and enzymatic properties of the catalytic domain of human complement protease C1r. *J Biol Chem* **276:** 36233–36240

Lacroix M, Rossi V, Gaboriaud C, Chevallier S, Jaquinod M, Thielens NM, Gagnon J, Arlaud GJ (1997) Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modelling. *Biochemistry* **36:** 6270–6282

Lamzin VS, Wilson KS (1997) Automated refinement for protein crystallography. *Methods Enzymol* **277:** 269–305

Le Y, Lee SH, Kon OL, Lu J (1998) Human-L-ficolin: plasma levels, sugar specificity, and assignment of its lectin activity to the fibrinogen-like (FBG) domain. *FEBS Lett* **425:** 367–370

Lim JH, Kim MS, Kim HE, Yano T, Oshima Y, Aggarwal K, Goldman WE, Silverman N, Kurata S, Oh BH (2006) Structural basis for the preferential recognition of diaminopimelic acid-type peptidoglycan by a subset of peptidoglycan recognition proteins. *J Biol Chem* **281:** 8286–8295

Liu Y, Endo Y, Iwaki D, Nakata M, Matsushita M, Wada I, Inoue K, Munakata M, Fujita T (2005) Human M-ficolin is a secretory protein that activates the lectin complement pathway. *J Immunol* **175:** 3150–3156

Lynch NJ, Roscher S, Hartung T, Morath S, Matsushita M, Maennel DN, Kuraya M, Fujita T, Schwaeble WJ (2004) L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement. *J Immunol* **172:** 1198–1202

Ma YG, Cho MH, Zhao M, Park JW, Matsushita M, Fujita M, Lee B (2004) Human mannose-binding lectin and L-ficolin function as specific pattern recognition proteins in the lectin pathway of complement. *J Biol Chem* **279:** 25307–25312

Matsushita M, Endo Y, Fujita T (2000) Cutting edge: complement-activating complex of ficolin and mannose-binding lectin-associated serine protease. *J Immunol* **164:** 2281–2284

Matsushita M, Endo Y, Taira S, Sato Y, Fujita T, Ichikawa N, Nakata M, Mizuochi T (1996) A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin. *J Biol Chem* **271:** 2448–2454

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

Matsushita M, Fujita T (2001) Ficolins and the lectin complement pathway. *Immunol Rev* **180:** 78–85

Matsushita M, Kuraya M, Hamasaki N, Tsujimura M, Shiraki H, Fujita T (2002) Activation of the lectin complement pathway by H-ficolin (Hakata antigen). *J Immunol* **168:** 3502–3506

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53:** 247–255

Navaza J (2001) Implementation of molecular replacement in AMoRe. *Acta Crystallogr D* **57:** 1367–1372

Nicholls A, Sharp KA, Honig B (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11:** 281–296

Orlando KA, Pittman RN (2006) Rho kinase regulates phagocytosis, surface expression of GlcNAc, and Golgi fragmentation of apoptotic PC12 cells. *Exp Cell Res* **312:** 3298–32311

Pelletier I, Hashidate T, Urashima T, Nishi N, Nakamura T, Futai M, Arata Y, Kasai K, Hirashima M, Hirabayashi J, Sato S (2003) Specific recognition of *Leishmania major* poly-beta-galactosyl epitopes by galectin-9: possible implication of galectin-9 in interaction between L. major and host cells. *J Biol Chem* **278:** 22223–22230

Ramasamy R, Reese RT (1986) Terminal galactose residues and the antigenicity of *Plasmodium falciparum* glycoproteins. *Mol Biochem Parasitol* **19:** 91–101

Rossi V, Bally I, Thielens NM, Esser AF, Arlaud GJ (1998) Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J Biol Chem* **273:** 1232–1239

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

Shrive AK, Tharia HA, Strong P, Kishore U, Burns I, Rizkallah PJ, Reid KB, Greenhough TJ (2003) High-resolution structural insights into ligand binding and immune cell recognition by human lung surfactant protein D. *J Mol Biol* **331:** 509–523

Sugimoto R, Yae Y, Akaiwa M, Kitajima S, Shibata Y, Sato H, Hirata J, Okochi K, Izuhara K, Hamasaki N (1998) Cloning and characterization of the hakata antigen, a member of the ficolin/opsonin p35 lectin family. *J Biol Chem* **273:** 20721–20727

Takahashi K, Ip WKE, Michelow IC, Ezekowitz RAB (2005) The mannose-binding lectin: a prototypic pattern recognition molecule. *Curr Opin Immunol* **17:** 1–8

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

Thielens NM, Cseh S, Thiel S, Vorup-Jensen T, Rossi V, Jensenius JC, Arlaud GJ (2001) Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL. *J Immunol* **166:** 5068–5077

Tsujimura M, Ishida C, Sagara Y, Miyazaki T, Murakami K, Shiraki H, Okochi K, Maeda Y (2001) Detection of a serum thermolabile β-2 macroglycoprotein (hakata antigen) by enzyme-linked immunosorbent assay using polysaccharide produced by *Aerococcus viridans*. *Clin Diagn Lab Immunol* **8:** 454–459

Turner MW (1996) Mannose-binding lectin: the pluripotent molecule of the innate immune system. *Immunol Today* **17:** 532–540

Wang X, Rocheleau TA, Fuchs JF, Christensen BM (2006) Beta 1, 3-glucan recognition protein from the mosquito, *Armigeres subalbatus*, is involved in the recognition of distinct types of bacteria in innate immune responses. *Cell Microbiol* **8:** 1581–1590

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

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

Yee VC, Pratt KP, Cote HC, Trong IL, Chung DW, Davie EW, Stenkamp RE, Teller DC (1997) Crystal structure of a 30 kDa C-terminal fragment from the gamma chain of human fibrinogen. *Structure* **15:** 125–138

Yuita H, Tsuiji M, Tajika Y, Matsumoto Y, Hirano K, Suzuki N, Irimura T (2005) Retardation of removal of radiation-induced apoptotic cells in developing neural tubes in macrophages galactose-type C-type lectin-1-deficient mouse embryos. *Glycobiology* **15:** 1368–1375

Zundel S, Cseh S, Lacroix M, Dahl MR, Matsushita M, Andrieu JP, Schwaeble WJ, Fujita T, Arlaud GJ, Thielens NM (2004) Characterization of recombinant mannan-binding lectin-associated serine protease (MASP)-3 suggests an activation mechanism different from that of MASP-1 and MASP-2. *J Immunol* **172:** 4342–4350