Immunobiology (2016.10) 221(10):1058-1067.

The collectins CL-L1, CL-K1 and CL-P1, and their roles in complement and innate immunity

Soren W.K. Hansena, Katsuki Ohtani, Nitai Roy, Nobutaka Wakamiya

Review

The collectins CL-L1, CL-K1 and CL-P1, and their roles in complement and innate immunity

Soren W. K. Hansen[a,1,*], Katsuki Ohtani[b,1], Nitai Roy[b] and Nobutaka Wakamiya[b,**]

[a]Department of Cancer and Inflammation Research, Institute of Molecular Medicine, University of Southern Denmark, Odense C, Denmark

[b]Department of Microbiology and Immunochemistry, Asahikawa Medical University, Asahikawa, Japan

*Corresponding author: Department of Cancer and Inflammation Research, Institute of Molecular Medicine, University of Southern Denmark, J.B. Winsløws Vej 21, 5000 Odense C, Denmark. Tel. +45 6550 3062.
E-mail address: shansen@health.sdu.dk (S.W. K. Hansen).
**Corresponding author: Department of Microbiology and Immunochemistry, Asahikawa Medical University, 2-1-1-1 Midorigaoka-Higashi, Asahikawa 078-8510, Japan, Tel.: +81 166 68 2393; fax: + 81 166 68 2399.
E-mail address: wakamiya@ashaikawa-med.ac.jp (N. Wakamiya).

[1]These two authors contributed equally and should be regarded as joint first authors.

Keywords: collectin, complement, MASP, collagen, scavenger receptor, microbes, 3MC syndrome, innate immunity, embryogenesis

Abstract

Both the complement system and collectins play important roles in our innate immune system. The collectins, which are characterized by their inclusion of a collagen-like region and a calcium-dependent carbohydrate recognition domain, are pattern recognition molecules and include the well characterized proteins mannan-binding lectin (MBL) and the surfactant proteins SP-A / -D. Collectin liver 1 (CL-L1), collectin kidney 1 (CL-K1) and collectin placenta 1 (CL-P1) are the most recently discovered collectins. Although their function is still under investigation, accumulating information suggests that CL-L1, CL-K1 and CL-P1 play important roles in host defense by recognizing a variety of microorganisms and interacting with effector proteins, including complement components. The recent establishment of the existence of CL-K1 in the circulation in form of heteromeric complexes with CL-L1 (known as CL-LK) and its activation of the lectin pathway via MASPs, drew new attention in the complement biology, which was further strengthened by the observed interactions between CL-P1 and CRP-C1q-factor H or properdin. Deficiency of either CL-K1 or MASP-3 has been demonstrated in 3MC syndrome patients with developmental abnormalities, showing that lectin pathway components, regulation and / or activation are essential during the embryonic development; another feature that they most likely share CL-P1.

Herein, we discuss the recent characteristics and roles of the collectins CL-L1, CL-K1 and CL-P1 in the complement system, in innate immunity and their possible association with disease development and pathogenesis.

Introduction

Bovine conglutinin was the first identified collectin and described as a vertebrate lectin, mediating agglutination of erythrocytes coated with antibody and complement (Bordet et al. 1906). Later it was found that conglutinin binds to both yeast cell walls and to the inactivated form of C3b, iC3b (Hirani et al. 1985). Lachmann demonstrated further that conglutinin included a collagen-like region (Davis AE 3rd et al. 1984). Seventy years after the discovery of conglutinin, the structurally related mannan-binding lectin was identified and characterized as a complement associated protein by the groups of Kawasaki and Kawakami. With the discovery of additional structurally similar

proteins, the term “collectin” was introduced in 1993 when three independent groups proposed to gather structurally related proteins into a distinct new group of proteins, the collectins, defined by being composed of both a collagen-like region and a C-type lectin domain (Sastry et al. 1993; Reid et al. 1993; Holmskov et al. 1993). This review focuses on the most recently described collectins: CL-P1 (collectin placenta 1), CL-L1 (collectin liver 1) and CL-K1 (collectin kidney 1), which were all identified between 1999-2006 by EST-screenings of collectin homologies and subsequently characterized by full-length cDNA cloning by Wakamiya and colleagues, and named accordingly to the tissue wherein the highest expression was observed (Ohtani et al. 2001; Ohtani et al. 1999; and Keshi et al. 2006). Their cDNAs encoding CL-P1, CL-L1 and CL-K1 were registered to GenBank under the following updated accession numbers NM_130386.2 (*COLEC12*), NM_006438.3 (*COLEC10*) and NM_024027.4 (*COLEC11*).

Due to a non-classical signal peptide, CL-L1 was in the originally work proposed to be an intracellular hepatic protein with weak affinity for mannose but later studies have in contrast shown that CL-L1 is also secreted and found in the circulation (see below) (Ohtani et al. 1999). CL-L1 and CL-K1 are referred as collectin 10 and collectin 11.

CL-P1 was in parallel with Wakamiya’s work characterized by Nakamura and colleagues, who identified CL-P1 as a scavenger receptor with C-type lectin domain (SRCL) (Nakamura et al. 2001). In the originally work, CL-P1 was reported to bind to both oxidized LDL and to various microorganisms, suggesting that it, not only due to its structure but also due to its functionality, bridged between innate immunity and scavenging of host molecules (Ohtani et al. 2001; Nakamura et al. 2001). Albeit there is no doubt that CL-P1 fulfills the requirements for being classified as a collectin, it is worth emphasizing that CL-P1 bears higher structurally resemblance and phylogenetic relationship with scavenger receptors and galactose binding C-type lectins, like the asialoglycoprotein receptor, than with the other collectins.

Among all the collectins in humans, CL-K1 is the most recently discovered collectin and was in the original work found in the circulation with the ability to bind to mannose- and fucose-containing microbial derived products, and hence proposed to play a role in humoral innate immunity (Keshi et al. 2006).

With the discovery of CL-P1, CL-L1 and CL-K1 and the recent establishment of heteromeric complexes of CL-L1 and CL-K1 (CL-LK), and their potent complement activating ability (Henriksen et al. 2013), additional aspects have been added to the growing intersection between collectin biology and the complement system. This intersection is further strengthened by the recent

findings of CL-P1 interacting with the complement components CRP/C1q and properdin (Roy et al. 2016; Ma et al. 2015). This review focuses on the biology of the collectins CL-P1, CL-L1 and CL-K1 and their association with the complement system (Table 1).

CL-L1 and CL-K1: Genes, transcription, tissue distribution and polymorphisms

The genes encoding human CL-L1 and CL-K1 are named *COLEC10* (Gene ID 10584) and *COLEC11* (Gene ID 78989) and located at chromosomal positions 8q23-24.1 (39,784bp) and 2p25.3 (49,398bp), respectively. In accordance with other collectin genes, the two genes are mosaics with separate exons encoding the carbohydrate recognition domain, the alpha helical coiled-coil region, the collagen-like region (up to three exons) and an initial combinatorial exon encoding the N-terminal, the signal peptide and some 5'-untranslated sequence. In the *COLEC11* gene there exist variations in the number and locations of upstream 5'-untranslated exons and hence the start location of transcription (Selman 2012). Neither of the two genes is located in the collectin locus found on human chromosome 10, encoding MBL, SP-A1/A2 and SP-D, but their compositions show clear evidence of evolutionary mutual gene duplication of an ancestor gene encoding CL-L1 / CL-K1 and further gene duplication into the ancestor gene(s) of MBL, SP-A and SP-D genes.

COLEC10 is transcribed and matured into a mRNA transcript of 1708 nucleotides, among which 834 nucleotides translate into a CL-L1 protein of 277 a.a., including a signal peptide of 27 a.a.. At present only one single mRNA transcript, representing full-length CL-L1, have been identified. This contrasts with the transcription of *COLEC11,* for which, at present, more than 11 isoforms of mRNA transcripts have been identified. In comparison with the full-length isoform, referred to as isoform a, some of the isoforms lacks wholly or partly exons encoding the collagen-like region (isoform c, d, e) whereas other differ in their start of transcription and inclusion of untranslated or alternative exons (b, f, g, h, i, j). However, the situation becomes simpler when looking at what have been detected at the protein level in the circulation. It appears that only full-length isoform a and isoform d are detected as proteins in the circulation (Henriksen et al. 2013). Due to the absence of a single collagen encoding exon, isoform d is characterized by a short collagen-like region of 42 a.a. vs. 66 a.a. in full-length CL-K1. It is worth noting, that full-length CL-K1 may also arise from isoform f with an alternative transcription start, which potentially could be transcriptionally regulated different from isoform a (Selman et al. 2012). Full-length CL-K1 derives from an mRNA transcript of 1304 nucleotides, among which 813 nucleotides translate into a protein of 271 a.a., including a signal peptide of 25 a.a..

mRNA transcript encoding CL-K1 and CL-L1 have been detected in various human tissues with the liver and the adrenal glands as organs with relative high amount of transcripts of both *COLEC11* and *COLEC10* (Hansen et al 2010; Ohtani et al. 1999).

On the protein level CL-K1 has by immunohistochemistry been localized within many tissues, with the adrenals, kidneys and liver as sites with the highest abundance of CL-K1. In the adrenals all three layers, zona fasciculate, glomerulosa and reticularis were associated with CL-K1 immunoreactivity. In the liver, hepatocytes appear as the cellular source of CL-K1 and in the kidneys, CL-K1 immunoreactivity was associated with particular the distal tubules but also with the glomerulus and proximal tubules (Hansen et al. 2010).

Only limited localization studies on CL-L1 have been carried out and showed that CL-L1 in agreement with its name is present in the liver and that the cellular source is hepatocytes (Ohtani et al. 1999).

Thus, at present hepatocytes in the liver seem to be the source of both CL-L1 and CL-K1 and the heteromeric form, CL-LK, found in the circulation. Further localization studies on especially CL-L1 are needed to elucidate overlapping sites of expression but based on the expression of mRNA the adrenals are a potential additional source of both CL-L1 and CL-K1. It is also worth noting that judged from the studies on mRNA, it appears that only CL-K1, and not CL-L1, is expressed in the kidneys. A detailed expressional and immunohistochemical study on mouse CL-K1 have been conducted and show good agreement between humans and mice (Motomura et al. 2008). A detailed comparison of expression and immunolocalization between CL-L1 and CL-K1 in both humans and mice has recently been reviewed elsewhere (Selman et al. 2012).

CL-K1 and CL-L1 have both been detected in the circulation and different ELISAs have subsequently been made allowing for their measurements (Keshi et al. 2006, Hansen et al. 2010; Yoshizaki et al. 2012; Selman et al. 2012; Axelgaard et al 2013). The average serum concentration of CL-K1 in healthy populations in Japan and Denmark was estimated to 340 ng/ml and 284 ng/ml, respectively (Yoshizaki et al. 2012; Selman et al. 2012). The average serum concentration of CL-L1 was initially overestimated but recently with the development of better standards by the same group found to be 306 ng/ml in a healthy Danish population (Axelgaard et al. 2013; Troldborg et al. 2015)..

Recently, Bayarri-Olmos and colleagues associated serum concentrations of CL-L1 and CL-K1 among healthy Danish Caucasians with polymorphisms in their respective genes (Bayarri-Olmos et al. 2015). Three non-synonymous variations were identified: CL-L1 p.Glu78Asp (rs150828850, collagen-encoding, minor allele frequency 0.003), CL-L1 p.Arg125Trp (rs149331285, neck-encoding, minor allele frequency 0.007) and CL-K1 p.His219Arg (rs7567833, carbohydrate recognition domain (CRD) -encoding, minor allele frequency 0.033). Only CL-L1 p.Arg125Trp appeared to have an effect on the serum concentration. Unfortunately, no homozygous for this polymorphism was identified but among the three heterozygous identified, the CL-L1 serum concentration was increased by approximately 40% (P = 0.0478).

A polymorphism in the promoter region of *COLEC11*, *COLEC11*-9570C>T (rs3820897) was identified and found to be associated with increased serum levels (at least 10%) among the three homozygous individuals tested (P=0.044). A minor but similar (insignificant) trend was observed among 41 heterozygotes.

In another study among Nigerian individuals of Yoruba ethnicity, the CL-K1 p.His219Arg (rs7567833) polymorphisms was the major allele, and associated with decreased CL-K1 serum levels, with approximately a 40% decrease in the median concentration between homozygotes, (P = 0.03, n=41) (Antony et al. 2015).

Very rare polymorphisms in the *COLEC11* gene encoding CL-K1 have been associated with the developmental syndrome known as 3MC (also discussed below). These encompass a total of six known mutations, with three resulting in substitutions or deletion of amino acid residues in the CRD (p.Ser169Pro , rs387907075,; p.Gly204Ser, rs387907076; and p.Ser217del, c.648-659delCTC,) and the other three in frame shift or lack of transcription (p.Phe16SerfsX85, c.56delC; p.Gly101ValfsX113, c.300delT; and a Exon 1-3 deletion) (Rooryck et al. 2011). Homozygosity for any of the six 3MC associated mutations result in the development of the syndrome and lack of CL-K1 in the circulation, while heterozygotes are healthy with approximately half of the ”normal” CL-K1 level (Rooryck et al. 2011; Selman et al. 2012).

Structural composition of CL-L1, CL-K1 and CL-LK

The primary structures of CL-L1 and CL-K1 follow that of the classical collectins (e.q. MBL but not CL-P1). This classical structure includes an N-terminal segment containing cysteine residues participating in cross-linking of polypeptide chains into subunits (made of three polypeptide chains) and further into oligomers of subunits. The N-terminal segment is followed by a collagen-like region, an alpha-helical coiled neck region and a C-terminally located carbohydrate recognition domain (CRD). In contrast with other collectins, both CL-L1 and CL-K1 include a C-X-C motif in the beginning of the neck region. Their CRDs are, in parallel with the CRDs of the classical collectins, a "short" form of C-type lectin domain with 4 cysteine residues, in comparison with the long form of C-type lectin domains (e.q. found in CL-P1), which includes additional residues and a total of 6 cysteines residues. CL-L1 possesses two N-linked glycosylation motifs in the CRD, in contrast to CL-K1, which lacks such motifs.

Overall there is a striking primary structural homology between CL-L1 and CL-K1, with identical numbers a.a. residues in every region, with the exception of the N-terminal segments and their signal peptides. Also, the locations of each of their 7 cysteine residues are identical.

Both CL-K1 and CL-L1 have been expressed in mammalian cells. However, their recombinant products are inferior in terms of oligomeric sizes and complement activity (discussed below), in comparison with that observed in the circulation (Keshi et al. 2006; Axelgaard et al. 2013; Hansen et al. 2010).

By means of a non-denaturating calcium sensitive anti-CL-K1 affinity chromatography both natural CL-K1 and CL-L1 were purified from plasma (Henriksen et al. 2014). Their characterization showed that they circulate in the form of oligomers made of heteromeric subunit complexes of CL-L1 and CL-K1, referred to as CL-LK (Henriksen et al. 2013) (Figure 1). By quantitative mass spectrometry the ratio of the two were approximately 2:1 (CL-K1:CL-L1) in a subunit, and by collagenase treatment it was further found that the C-X-C motif in the neck region participates in crosslinking of CL-L1 and CL-K1 polypeptides within a subunit (Henriksen et al. 2013 and unpublished S.W.K. Hansen). From the purified plasma preparation of CL-L1 and CL-K1 it appears that the majority of CL-L1 and CL-K1 in the circulation are in the form of heteromeric complexes. Estimation of CL-L1 serum levels in CL-K1 deficient persons with the 3MC syndrome showed that the patients have only 10-15% of the "normal" serum level of CL-L1, suggesting that the presence of CL-K1 and heteromer formation is required for optimal synthesis, stability or secretion into

circulation (unpublished Steffen Thiel Aarhus University and S.W.K. Hansen). To further support the notion that the majority of CL-L1 and CL-K1 in the circulation is in the form CL-LK, there is a nearly complete correlation between their respective serum levels among healthy individuals, meaning that a high CL-K1 serum concentration associate with a high CL-L1 serum concentration, and vice versa (Troldborg et al 2015; Bayarri-Olmos 2015).

Although CL-LK is the first heteromeric collectin characterized, heteromerity is also known from other soluble collagen-like molecules, i.e. C1q, consisting of three different polypeptides (C1qABC), and also from other molecules with alpha-helical coiled-coils, i.e. the matrilins, which also have the C-X-C motif (Reid et al. 1976; Frank et al. 2002).

As the purification of CL-LK favorized enrichment of small oligomers, the size distribution of natural CL-LK oligomers awaits a precise characterization. However, judged from gel permeations chromatographies and Western blotting analyses of serum and plasma, it appears that CL-LK, like MBL, are heteromeric in terms of oligomeric sizes, ranging from as small as a single subunit to at least hexamers of subunits (Axelgaard et al. 2013; Henriksen et al. 2013; Hansen et al. 2010).

Recently, a recombinant fragment of CL-K1 representing the neck and CRD was expressed in *E.coli*, solubilized from inclusion bodies and crystalized to a resolution of 2.45 Å, with and without carbohydrate ligands (discussed further below). In parallel with the structures of other collectins, the structure was trimeric and stabilized by non-covalent inter-polypeptide chain interactions in the alpha-helical coiled-coil region. The C-X-C motif was not included in the crystalized fragment (Venkatraman et al. 2015) (Figure 1).

CL-L1 and CL-K1: Ligand specificity, binding to microbial ligands – and inhibition hereof

It is important to emphasize that many of the following cited studies have been carried out in "clean systems" in absence of serum/plasma components and that it is known that serum/plasma components interfere heavily with the interaction between at CL-K1 / CL-L1 and carbohydrates (Henriksen et al. 2013). Using traditional ELISAs with immobilized carbohydrate ligands results in no detection of CL-K1 or CL-L1 binding to the wells in the presence of serum or plasma incubated with carbohydrate conjugated beads result in only a relative small fraction CL-L1 and CL-K1 binding to the beads (Henriksen et al. 2013; personalized correspondence Steffen Thiel, University of Aarhus, Denmark). It is most likely a matter of equilibrium between low vs. high affinity ligands and this aspect is discussed further at the end of this section.

During the initial characterization of CL-L1 a recombinant fragment, representing the CRD conjugated to various tags, was used in both ELISAs and in sugar blot analysis with immobilized CL-L1 and probing with biotinylated-carbohydrates. CL-L1 showed calcium dependent lectin activity for D-mannose, *N*-acetylglucosamine, D-galactose and D-fucose (Ohtani et al. 1999). More recently Axelgaard and colleagues studied the CL-L1 monosaccharide specificity in an elegant setup using serum CL-L1 as their source, and found that CL-L1 bound to L-fucose, D-mannose, and *N*-acetylglucosamine (Axelgaard et al. 2013). However, this corresponds to the specificity of CL-K1, and as CL-L1 later was shown to circulate in the heteromeric complex CL-LK, it is likely that these results were greatly influenced by binding of CL-K1 to the found ligands. Looking at the central triad of a.a. residues in the CRD, responsible for galactose vs. mannose specificity, QPD vs. EPN, respectively, CL-L1 possesses an EPS motif. The EPS motif is known to provide some preference for mannose-like carbohydrates and is found in chicken SP-A (cLL) (Hogenkamp et al. 2008). The lectin activity of CL-L1 may be further complicated or even compromised by the presence of an *N*-linked glycosylation present on Arginine 252; an a.a. residue known from other collectins to be involved in both interactions with the bound monosaccharide and in coordination bonds with the calcium atom, which facilitate assembly of the binding site. Based on the above experiments it is evident that further studies with CL-L1 expressed in mammalian cells, fully or unglycosylated, are needed in order to characterize the specificity of CL-L1.

On the monosaccharide level CL-K1 has specificity for L-fucose, D-mannose and *N*-

acetylmannosamine (Ohtani et al. 1999; Hansen et al. 2010). Recently the specificity of CL-K1 was further analyzed by glycan array screenings and convincingly verified by successful co-crystallization of the neck-CRD fragment in complex with the identified ligands (Venkatraman et al. 2015). These studies showed that CL-K1 bound selectively and strongly to the disaccharide D-mannose (α1-2)-D-mannose located terminally and found in both non-self and self high mannose structures. Unusually, in comparison with other collectins, the binding took advantage of interaction with both mannose units of the disaccharide. The terminal non-reducing mannose, with a glycoside interaction at its position 1, which normally interacts with the central EPN motif in other collectins, interacted with a proximal site composed of an arginine (position 200) and a glutamate (position 244) residue, to which hydrogen bonds between its 2- and 3-OHs were formed. In combination with these interactions, the reducing mannose (penultimate in an oligosaccharide = second terminal), with a glycoside linkage at in its position 2, interacted with the central EPN motif and formed the classical coordination bonds with the calcium atom, via its equatorial 3- and 4-OHs (Venkatraman et al. 2015).

In addition to interaction with carbohydrates, CL-K1 has in parallel to MBL been shown to interact calcium independently with negatively charged ligands e.q DNA, RNA and sulfated compounds, i.e. heparin (Henriksen et al. 2013). Binding to negatively charged ligands took place at physiological pH and ionic strength, but was sensitive to hyper salinity. Interaction with surfaces coated with both carbohydrates and negatively charged ligands increased the binding synergistically, implying that optimal binding to surfaces included simultaneously interactions with both carbohydrates and negatively charged molecules. CL-K1 binding to apoptotic cells appeared also to be mediated via interaction with both carbohydrate and negatively charged ligands (Henriksen et al. 2013). CL-K1 has been shown to bind to the following microorganisms, or cell wall or membrane components, hereof: *E.coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Mycobacterium tuberculosis, Saccharomyces cerevisiae, Candida albicans,* and Influenza A virus (Keshi et al. 2006; Hansen et al 2010; Troegeler et al. 2015).

It is expectable that the overall specificity of CL-LK found in the circulation is contributed by both the specificity of CL-K1 and CL-L1. On the one hand, this may provide an increased range of “low” affinity ligands and on the other hand provide strict requirement of “high” avidity ligands that are recognized simultaneously by CL-K1 and CL-L1. It is likely that such requirements for - and equilibrium of - CL-LK ligands combined with both CL-K1’s binding to self and non-self high

mannose structures, and also to negatively charged ligands, may explain some of the observed inhibition of CL-LK binding in the presence of serum (introduced above). It appears that serum/plasma has to be diluted 1,000 – 10,000 times before binding of CL-LK at physiological relevant concentration takes place (Henriksen et al. 2014), most likely to prevent uncontrolled complement activation. A few reports have demonstrated binding / complement activity in the presence of serum but in every case, the concentration of CL-K1 were hyper physiological (10,000 – 1,000,000 times) or relevant controls in form of complete CL-K1 depletion from serum by ELISA verification, or CL-LK control in comparison with CL-K1, have been omitted (Venkatraman et al. 2015; Ma et al. 2013). Although the average serum levels of CL-L1 and CL-K1 are below that of MBL (0.3 µg/ml vs. 1.5 µg/ml), respectively, it is likely that the observed inhibition of CL-LK binding by serum components is a rational explanation why it fails to co-purify with MBL from serum by carbohydrate affinity chromatographies, and hence its only recent discovery.

CL-LK, activator of the lectin pathway of the complement system

The full potential of CL-L1 and CL-K1 to activate the complement system was established by the characterization of the CL-LK heteromer in the circulation (Henriksen et al. 2013). Unlike CL-L1 or CL-K1 alone, the CL-LK is a convincing activator of the lectin activation pathway of the complement system via MASP-2 and MASP-1, at the same level as that of MBL. It was evident that the complement activation by CL-LK depended on high oligomerity, in accordance with what recently have been described for MBL (Kjaer et al. 2016). Whether the dependency of high oligomerity was required to achieve high avidity interaction with the immobilized ligands or alternative to obtain strong interaction with particular MASP-2, remains to be established. In the circulation CL-LK was found in already formed complexes with MASP-1/3 and MASP-2, and in *in vitro* settings, CL-LK also bound to carbohydrate or negatively charged ligands, and then interacted with free MASP-2 to initiate activation of C4 (Henriksen et al. 2013) (Figure 2).

In preceding work, the interaction between CL-L1 or CL-K1 alone with MASP-1/3 was established but the interaction with MASP-2, and hence complement activation, was inferior in comparison with that of MBL (Hansen et al. 2010; Ma et al. 2013, Axelgaard et al. 2013). It is likely that the assembly of the CL-LK heteromer, in comparison with CL-L1 and CL-K1 alone, facilitate both optimal assembly of the collagen-like motif, whereto MASPs bind, and high oligomerity, in order to interact strongly with the MASPs, and in particular with MASP-2 (Henriksen et al. 2013).

CL-L1 and CL-K1: Association with diseases

In the work by Rooryck and colleagues, deficiency of either CL-K1 or MASP-3 was shown to cause the rare autosomal recessive syndrome known as 3MC (Mingarelli, Malpuech, Michels and Carnevale Syndrome), characterized by various symptoms including facial, genital, renal, mental, and limb abnormalities (Rooryck et al. 2011). The corresponding gene defects and variations were introduced above. 3MC occurs during embryonic development, and since defects of CL-K1 *(COLEC11*) and MASP-3 (*MASP1*) were independently associated, the work convincingly demonstrated a novel importance of complement components during embryogenesis. By studies in Zebra fish it was further suggested that CL-K1 served as a chemotactic attractant to guide migration of neural crest cells (Rooryck et al. 2011). To further support the above findings, it was recently demonstrated that developmental delay / intellectual disability was associated with hyper methylation of *COLEC11*, but eventual CL-K1 deficiency in the affected individuals was not investigated further, although a single patient had some overlapping symptoms with that of 3MC patients (Kolarova et al. 2015).

Similar with the function of other complement components, CL-K1 regulation and potentially also function have been associated with the coagulation system.
Takahashi and colleagues found that the incidence of elevated plasma levels of CL-K1 was significantly higher (median concentration increased by approximately 75%) in patients with disseminated intravascular coagulation (DIC), compared to non-DIC patients, with an odds ratio of 1.929 (confidence interval 1.041–3.866) (Takahashi et al. 2014). DIC results from dysregulation of the coagulation system that may occur during profound activation of the innate immune system, and leads to both thrombosis and bleeding at the same time. Independent of the development of DIC, high plasma levels of CL-K1 was associated with coagulation disorder (and respiratory disorders), and it was suggested that CL-K1 may serve both as a biomarker of the development of DIC and potentially also be a contributing factor to the DIC development (Takahashi et al. 2014).

Although the binding of CL-K1 to the digenetic trematode *Schistosoma haematobium* not has been shown directly, high serum levels of CL-K1 among Nigerian populations are associated with decreased appearance of urogenital infection with *Schistosoma haematobium* (Antony et al. 2015). The IL-6 level among infected individuals was further shown to correlate with the serum level of CL-K1 (Antony et al. 2015). It is worth emphasizing that neither CL-K1 nor CL-L1 are traditional

acute phase proteins. Their serum levels do not correlate with increased levels of traditional inflammatory mediators, including CRP and TNF-α (Axelgaard et al. 2013; N. Wakamiya and S.W.K. Hansen unpublished results).

CL-LK has also been shown to bind to *Mycobacteria tuberculosis* via interaction with the lipoglycan, lipoarabinomannan (ManLAM), which is major constituent of their cell envelope. The serum level of CL-K1 was further reduced in patients with tuberculosis, suggests either a consumption of CL-K1 or that low levels are associated with increased risk of infection or colonization (Troegeler et al. 2015).

In a cross-sectional cohort study of systemic lupus erythematosus (SLE), decreased serum levels of CL-L1 were associated with SLE. The median plasma concentration was decreased by approximately 20% ($P>0.001$) in SLE patients ($n=58$) in comparison with healthy ($n= 65$). Also here, it remains to be clarified whether the decreased levels were associated with polymorphisms or increased CL-L1 consumption (Troldborg et al. 2015).

CL-P1: Gene, structure and tissue distribution

Whereas all other collectins are soluble, CL-P1 is a single pass (type II) transmembrane protein orientated with its N-terminal towards the cytosol. The HUGO Gene Nomenclature Committee (HGNC) approved *COLEC12* as CL-P1 gene symbol (Gene ID 81035). Its cytogenetic location is 18p11.32 with the genomic coordinates (GRCh38) being 18:319,354-500,728 (NCBI). Its cDNA sequence revealed that *COLEC12* has an open reading frame of 2226 base pairs encoding a sequence of 742 amino acids (Ohtani et al. 2001). *COLEC12* cDNA was initially cloned from the placenta and the liver although the mRNA is expressed ubiquitously in many tissues. Northern blotting showed a major band of approximately 3.2 kilobases in various adult tissues; the highest expression was found in the placenta, the lung and the heart (Ohtani et al. 2001; Nakamura et al. 2001). Deduced amino acid sequence revealed that it contains an N-terminal intracytoplasmic domain, a transmembrane domain, a collagen-like region, a coiled-coil domain and a CRD localized C-terminally.

In humans, the full-length CL-P1 was named isoform I, whereas isoform II referred to a variant lacking the CRD (Ohtani et al. 2001). Recently, additional isoforms were characterized and some of these appear only as intracellular proteins (Mori et al. 2014).

The cytoplasmatic N-terminus contains an internalization signal YKRF involved in CL-P1-mediated endocytosis and signaling (discussed below). Although not verified, it is likely that the CL-P1 is organized as a trimer of polypeptide chains, initiated and stabilized by formation of the α-helical coiled-coil and the collagen-like region. It was found that CL-P1 has an approximate molecular mass of 140 kDa in both HUVEC and in placenta membrane extracts and a mass of 90 kDa as a deglycosylated size SDS-PAGE under reducing condition.

As stated in the introduction, CL-P1 may be regarded as both a collectin and a scavenger receptor; with the scavenger receptors SR-AI and SR-A3 being the closest related scavenger receptor homologies, and the asialoglycoprotein receptor and macrophage galactose lectin being the closest related C-type lectin homologues. The α-helical coiled-coil and collagen-like region are the *bona fida* characteristics of the scavenger receptors, and CL-P1 possesses three positive charged clusters in the collagen-like region, which are highly conserved in CL-P1 from various species (Mori et al. 2014).

Similar positive charged clusters from other scavenger receptors are known to be essential for the

interaction with negatively charged ligands. Other scavenger receptors possess only between none to two positive charged clusters; illustrating the potency of CL-P1 for interaction with negatively charged ligands.
During evolution, CL-P1 likely evolved by exon shuffling, and genes encoding CL-P1 have been identified in species ranging from bone fish to humans.

On the protein level CL-P1 was detected in *in vitro* cultured human umbilical vein endothelial cells (HUVEC), human umbilical artery endothelial cells (HUAEC) and arterial endothelial cells in mice, but not in monocyte-macrophage lineage cells (Ohtani et al. 2001). Northern blot analyses indicate that CL-P1 also is expressed by astrocytes, microglia, vascular / perivascular cells of the brain and in cultured nurse-like cell lines (Yoshida et al. 2003; Nakamura et al. 2006)
By traditional immunohistochemistry techniques using monoclonal antibodies it was suggested that CL-P1 was predominantly associated with the placenta and the retina (Selman et al. 2008). A minor expression of CL-P1 was found in mammary gland, the heart and the skeletal muscle. It was also suggested that CL-P1 mRNA synthesis / protein expression was heterogenous among cultured endothelial cells, and mainly associated with HUAECs, HUVECs, and the endothelial cell line HPMEC ST1.6R.

CL-P1-mediated endocytosis of glycans, microorganisms and OxLDL

In principal there are two pathways of endocytosis, either through clathrin-coated vesicles or via lipid rafts/caveolae. Ohtani and colleagues demonstrated that CHO cells expressing CL-P1 (CHO-CL-P1) bound to OxLDL but not to Ac-LDL or LDL (Ohtani et al. 2001). The binding of OxLDL to CHO-CL-P1 cells was inhibited by the negative polycharged molecules poly-I/G and dextran sulphate (Ohtani et al. 2001), suggesting that it is the positive charged clusters on the collagen-like region that involved in this interaction. Recently, Jang and colleagues found that YKRF sorting motif in cytoplasmic region of CL-P1 was associated with the μ2 chain of AP-2 adapter (Jang et al. 2014). They demonstrated that AP-2-μ2 was essential for CL-P1-mediated endocytosis of OxLDL, glycans and anti-CL-P1 antibodies, and that the endocytosis hereof depended on clathrin, dynamin and adaptin complex molecules. This indicated that CL-P1 is involved in the regulation of OxLDL in the circulation via clathrin-coated vesicles.

CL-P1 was found to bind to and mediate phagocytosis of *Saccharomyces*, *E. coli* and *S. aureus* (Ohtani et al. 1999; Jang et al. 2009). Furthermore, Jang and colleagues demonstrated the importance of CL-P1 in phagocytosis of zymosan in HUVECs. The use of small interfering RNAs of CL-P1 or other scavenger receptors confirmed that it was predominantly CL-P1, which mediated phagocytosis of zymosan, and only to a lesser extent other scavenger receptors. As an indicator of the physiological importance of this observation, it was also observed that serum actually increased the CL-P1-mediated phagocytosis several fold. It was demonstrated by deletion mutations that the positive charged clusters were essential for the interaction with both bacteria and OxLDL; with coiled-coil region contributing to optimal biding. At present it remains unknown why the coiled-coil region contributed to the binding. For the interaction with zymosan, it was in the same work shown, that both the CRD and the collagen-like region were involved (Mori et al. 2014).

CL-P1: specificity for carbohydrate ligands

The central a.a. triad of the CRD of CL-P1 is QPD, which is predicted to be specific galactose-type ligands, and is in agreement with the observation that CL-P1 bound to T-antigen and Tn-antigen in a Ca^{2+} dependent manner (Yoshida et al 2003). Tn antigen (GalNAcα1-O-Ser/Thr) is a universal carcinoma marker and is thought to participate in metastasis (Springer et al.1983; Springer et al.1984; Hirao et al. 1993).

By the use of glycan array analyses, Coombs and colleagues characterized that the CRD of human CL-P1 interacted with Lewis^x (Le^x) antigen and this observation was recently also confirmed by others (Coombs et al. 2005; Mori et al. 2014).

Le^X represents the trisaccharide Galβ1-4(Fucα1-3)GlcNAc. It appeared that the mouse CL-P1 has broader specificity, comprising interactions with both Le^X and Lewis A (Le^A), Galβ1-3(Fucα1-4) GlcNAc (Feinberg et al. 2007). In the same work the CRD of mouse CL-P1 was crystalized in complex with Le^X. The study demonstrated that the galactose residue in Le^X interacted with the QPD binding site within the CRD, with the equatorial 3- and axial 4-OH groups on galactose forming coordination and hydrogen bonds similar to those seen in other galactose-binding C-type CRDs (Feinberg et al. 2007) (Figure 3). This interaction orientated Le^X so that the terminal fucose residue contacted the CRD in the secondary binding site, whereas the central GlcNAc residue pointed away from the protein. It is worth noting that sialyation of Le^X and Le^A (into sLe^X and sLe^A), which occurs on the 3-hydroxyl group of galactose, is likely to affect negatively the interaction with the central binding and explain why CL-P1 does not bind to the majority of Lewis structures that are sialylated.

Le^X is also known as stage specific embryogenic antigen 1 (SSEA-1, CD15) and is a marker for undifferentiated cells including murine pluripotent stem cells, neutrophils, certain cells in patients with Hodgkin disease, some B-cell chronic lymphocytic leukemias, acute lymphoblastic leukemias, and most acute nonlymphocytic leukemias. In murine pluripotent stem cells Le^X plays an important role for the adhesion and migration of the cells in the initial embryogenesis.

Elola and colleagues showed the adhesion of polymorphonuclear granulocytes (PMNs) or human breast carcinoma cell, (MCF-7) to LPS-activated HUVECs, was mediated via Le^X on the PMNs or on human breast carcinoma cell and CL-P1 expressed by the HUVECs (Elola et al. 2007). The restricted glycan binding of CL-P1 and its distribution on vascular endothelial cells suggest that it might have a role in cell adhesion analogous to the leukocyte-endothelial interactions mediated by

selectins. In addition, the Le^x antigen is present on various tumors and could mediate the adhesion of cancer cells to vascular endothelium through CL-P1. This interaction might be used to prevent tumor metastasis.

Recently, Graham and colleagues demonstrated that CL-P1 in vascular endothelial cells bound and mediated endocytosis neutrophil granule glycoproteins (Graham et al. 2011). These glycoproteins possess a multi clustered terminal Le^x on a heterogenous mixture of branched glycans containing partially poly-N-acetyllactosamine residues. The abundant presence of Le^X in granule protein glycans can thus target granule proteins for clearance by CL-P1. These results suggest that Cl-P1 in the vascular spaces acts as a scavenger receptor to scavenge released neutrophil glycoproteins, both locally at sites of inflammation or systemically when they are released in the circulation.

CL-P1's role in complement activation and regulation

C-reactive protein (CRP) is the major acute phase protein in humans and expressed as part of the innate immune response to infection and tissue injury (Gabay et al. 1999). CRP and other pentraxins have been shown to activate a complement pathway through C1q, lectins, or on microbial surfaces. Du Clos and colleagues found that local production and activation of the complement play an important role in ischemic/reperfusion injury and allograft rejection (Du Clos et al. 2011). By localizing at sites of tissue damage, CRP has the potential to contribute to complement activation at these sites, but CRP may also regulate this activation by interacting with the complement factor H (CFH) bound to the surface of vascular endothelial cells. Very recently, Roy and colleagues demonstrated that CL-P1 bound CRP in a charge-dependent manner. This interaction mediated both activation of the classical complement activation pathway through C1q and drove an alternative amplification pathway using properdin, in the presence of (human) serum (Roy et al. 2016) (Figure 4). However, in the same setting, CRP also recruited CFH to the cell surfaces of CL-P1 expressing cells, and protected the formation of a terminal complement complex on normal healthy cells in the presence of serum. As deficiency or dysfunction of factor H is known to be involved in various autoimmune diseases, the recent observations of CL-P1-mediated complement activation / regulation may provide novel insight to the complement related pathology of autoimmune diseases and potentially become a target of therapy.

By ELISA measurements Ma and colleagues recently found a soluble form of CL-P1 in umbilical cord plasma with a median concentration of approximately 90 ng/ml; however, the form was undetectable in venous blood from healthy adults (Ma et al. 2015). It was speculated that this fetal and soluble form represented a shed form of CL-P1, in parallel with that of other scavenger receptors, i.e. CD163. When recombinant CL-P1 was expressed, they showed that soluble CL-P1 interacted with properdin, the only positive regulator of the complement system, which acts both as an initiator and stabilizer of the alternative pathway. Furthermore, they found that soluble CL-P1 recognized *Aspergillus fumigatus* partially through the CRD in a Ca^{2+} independent manner and, via interaction with properdin, lead to activation with the alternative pathway (Figure 4). However, they found also that the CL-P1-properdin mediated complement activation fail to apply to the membrane-localized form observed in adult. It was speculated that the CL-P1-properin mediated complement activation may contribute adjunct to antibiotic therapy in the treatment of microbial

infections (Ma et al. 2015).

CL-P1 in disease and developmental process

Two unique findings have been reported in animal models. Koyama and colleagues recently demonstrated the induction of CL-P1 under *in vivo* and *in vitro* conditions (Koyama et al. 2011). Hypoxia/reoxygenation stimulation induced CL-P1 mRNA and protein expression in HUVECs. Furthermore, the exposure of ischemic / reperfusion upregulated CL-P1 mRNA and protein in a rat model. This pattern of inducible expression of CL-P1 *in vitro* and *in vivo*, due to oxidative stress, provides new insights in the development of endothelial dysfunction and the pathogenesis of atherosclerosis. Fukuda and colleagues cloned the cDNA of a zebrafish CL-P1. Their further findings in zebrafish embryo models showed that knockdown of zebrafish CL-P1 expression caused severe defects in vasculogenesis and development during early embryogenic stages (Fukuda et al. 2011). The knockdown experiment in zebrafish CL-P1 suggests a pivotal role for CL-P1 in fundamental developmental processes, although the detailed mechanism remains to be determined.
The first association between CL-P1 and disease was observed by Nakamura, who showed that CL-P1 was expressed in glial and vascular / perivascular cells from patients with Alzheimer's disease but not in cells from healthy persons (Nakamura et al. 2006). They suggested that CL-P1 might participate in clearance (scavenging) of amyloid beta protein. They further demonstrated the same findings in a mouse model for Alzheimer's disease. Recently, the genome-wide association studies (GWAS) have become popular, since the next generation sequencing machines has been convenient in nearly every country. The *COLEC12* gene, encoding CL-P1, has been found as a candidate gene for rheumatoid arthritis by the first GWAS (Srivastava et al. 2009). Two other GWAS described CL-P1 association with breast and thyroid cancer and third one with diabetic retinopathy (Peng et al. 2015; Espinal-Enriquez et al. 2015; von der Heyde et al, 2015). Danfeng Peng and colleagues demonstrated by GWAS that polymorphisms in or near the genes *ZNRF1*, *COLEC12, SCYL1BP1,* and *API5* were associated with diabetic retinopathy in Chinese patients with type 2 diabetes. (Peng et al. 2015). This study was performed by using 1,972 Chinese diabetes patients including 789 diabetic retinopathy patients. It showed that rs599019, located near to *COLEC12*, was associated with diabetic retinopathy (OR 0.835, p=0.0116) and with severity of diabetic retinopathy (p=0.0252.). The GWAS of three subtypes of thyroid carcinoma were carried out to identify significant differences in expression between the three subtypes papillary, follicular, anaplastic thyroid cancers (Espinal-Enriquez et al. 2015). They demonstrated that *COLEC12* gene products (mRNA) were up-regulated in papillary and follicular thyroid cancer in comparison to anaplastic

thyroid cancer, while the *GPR110* gene could be tested to distinguish papillary thyroid cancer over other tumor subtypes. They also suggested that *COLEC12* gene products are a possible candidate biomarker of anaplastic thyroid cancer. Their GWAS demonstrated the preponderance of matrix metalloproteinases pathway-dysregulation mechanisms over simple gene-malfunction as a main mechanism involved in the development of cancer phenotype. Another GWAS with *COLEC12* was shown in the detection of genes and SNPs affecting trastuzumab efficiency in breast cancer cell culture (von der Heyde, 2015). These authors demonstrated an increased expression of *COLEC12* mRNA in BT474 breast cancer cell, as model of trastuzumab sensitivity, and a decreased *COLEC12* mRNA level in breast cancer cells HCC 1954, as model intrinsic trastuzumab resistant cell. Another microarray study indicated that the over expression of *COLEC12* gene is found in BT474 cells and that *COLEC12* gene is likely to be an estrogen receptor (ER) primary target gene. These recent genome-wide association study and genome-wide expression analysis indicate that CL-P1 might have some role in the pathology and potentially also on the onset of several diseases.

Conclusions and future perspectives

Already by Bordet's characterization of the first collectin, bovine conglutinin, and its interaction with complement, in the form of the iC3b molecule, there appeared to be a tie between the two participants in innate immunity, the collectins and the complement system. This tie was made stronger by the discovery of MBL and its complement activation ability, and made even stronger by the similar complement activating potential of CL-LK and CL-P1's complement activating /regulating role, which we have reviewed above. It appears as if the functions of collectins and complement are intertwined for mutual benefits, in order to provide the strongest and most appropriate innate immune response. We have still not completely understood the biological roles of CL-LK and CL-P1 in relation with the complement system, and future studies should aim at addressing such aspects; seeking out under which disease associated circumstances and for which microorganisms these collectins prove to be essential.

Another striking overlap of function is that both CL-K1 and CL-P1 appear to play roles in the embryonic development. CL-K1 deficiency results in the developmental syndrome 3MC in humans and CL-P1 deficiency results in defects both the vasculogenesis and embryogenesis in the zebra fish (reviewed above). For both observations we need to fully understand the molecular mechanisms leading to these defects, and future studies should address such areas of interest, in addition to the roles of CL-L1 and complement in the same and similar defects. Although the complement associated role of MASP-3 is not as well established as that of MASP-1/-2, based on the observation that either CL-K1 or MASP-3 deficiency lead independently to the 3MC syndrome, it is likely that complement activation or regulation hereof is essential for embryonic development.

In the light of the reviewed association studies for levels, polymorphisms or regulation of CL-L1/-K1/-P1, it is also for future studies rational to address their roles in diseases such as thrombotic disorders, SLE, rheumatoid arthritis, breast cancer, thyroid cancer and diabetic retinopathy.

Due to the recently observed association between these three collectins and complement, it appears that we are emerging into a new sub field of collectin / complement research, which might hold promises for better understanding of the pathology of several diseases in addition to embryonic processes.

Acknowledgements

This work (N.W.) was supported by grants from JSPS KAKENHI Grant Number (22390113, 26293124), the Northern Advancement Center for Science & Technology (Sapporo, Japan), the Smoking Research Foundation and the Mizutani Foundation for Glycoscience. This work (S.W.K.H) was further supported by grants from A. P. Moeller Foundation, the Augustinus Foundation, the Family of Hede Nielsen Foundation, the Dagmar Marshall Foundation and the Højbjerg Foundation.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

Antony, J.S., Ojurongbe, O., Kremsner, P.G., Velavan, T.P., 2015. Lectin complement protein Collectin 11 (CL-K1) and susceptibility to urinary schistosomiasis. PLoS Negl. Trop. Dis. 25, e0003647.

Antony, J.S., Ojurongbe, O., Meyer, C.G., Thangaraj, K., Mishra, A., Kremsner, P.G., Velavan, T.P., 2015. Correlation of Interleukin-6 levels and lectins during Schistosoma haematobium infection. Cytokine 76, 152-155.

Axelgaard, E., Jensen, L., Dyrlund, T.F., Nielsen, H.J., Enghild, J.J., Thiel, S., Jensenius, J.C., 2013. Investigations on collectin liver 1. J. Biol. Chem. 288, 23407-23420.

Bayarri-Olmos, R., Hansen, S., Henriksen, M.L., Storm, L., Thiel, S., Garred, P., Munthe-Fog, L., 2015. Genetic variation of COLEC10 and COLEC11 and association with serum levels of collectin liver 1 (CL-L1) and collectin kidney 1 (CL-K1). Plos ONE 10, e0114883.

Bordet, J., Gay, F.P., 1906. Sur les relations des sensibilisatrices avec l'alexine. Ann. Imt. Pasteur 20, 467-498.

Coombs, P.J., Graham, S.A., Drickamer, K., Taylor, M.E., 2005. Selective binding of the scavenger receptor C-type lectin to Lewisx trisaccharide and related glycan ligands. The J. Biol. Chem. 280, 22993–22999.

Davis, AE 3rd., Harrison, R.A., Lachmann, P.J., 1984. Physiologic inactivation of fluid phase C3b: isolation and structural analysis of C3c, C3d,g (alpha 2D), and C3g. J. Immunol. 132, 1960-1966.

Du Clos, T.W., Mold, C., 2011. Pentraxins (CRP, SAP) in the process of complement activation and clearance of apoptotic bodies through Fcγ receptors. Curr. Opin. Organ Transplant. 16, 15-20.

Elola, M.T., Capurro, M.I., Barrio, M.M., Coombs, P.J., Taylor, M.E, Drickamer, K., Mordoh, J., 2007. Lewis x antigen mediates adhesion of human breast carcinoma cells to activated endothelium. Possible involvement of the endothelial scavenger receptor C-type lectin. Breast Cancer Res. Treat. 101, 161–174.

Espinal-Enríquez, J., Muñoz-Montero, S., Imaz-Rosshandler, I., Huerta-Verde, A., Mejía, C., Hernández-Lemus, E., 2015. Genome-wide expression analysis suggests a crucial role of dysregulation of matrix metalloproteinases pathway in undifferentiated thyroid carcinoma. BMC Genomics 16, 1-22.

Feinberg, H., Taylor, M.E, Weis, W.I., 2007. Scavenger receptor C-type lectin binds to the leukocyte cell surface glycan Lewis(x) by a novel mechanism. J. Biol. Chem. 282:17250-17258.

Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Mini, T., Jeno, P., Engel, J. Kammerer, R. A., 2002. b Characterization of the matrilin coiled-coil domains reveals seven novel isoforms. J. Biol. Chem. 277, 19071-19079.

Fukuda, M., Ohtani, K., Jang, S.J., Yoshizaki, T., Mori, K., Motomura, W., Yoshida, I., Suzuki, Y., Kohgo, Y., Wakamiya, N., 2011. Molecular cloning and functional analysis of scavenger receptor zebrafish CL-P1. Biochim. Biophys. Acta 1810, 1150–1159.

Gabay, C., Kushner, I., 1999. Acute-phase proteins and other systemic responses to inflammation. N. Engl. J. Med. 340, 448 –454.

Graham, S.A., Antonopoulos, A., Hitchen, P.G., Haslam, S.M., Dell, A., Drickamer, K., Taylor, M.E., 2011. Identification of neutrophil granule glycoproteins as Lewis x containing ligands cleared by the scavenger receptor C-type lectin. J. Biol. Chem. 286, 24336–24349.

Hansen, S., Selman, L., Palaniyar, N., Ziegler, K., Brandt, J., Kliem, A., Jonasson, M., Skjoedt, M.O., Nielsen, O., Hartshorn, K., Jørgensen, T.J., Skjødt, K., Holmskov, U., 2010. Collectin 11 (CL-11, CL-K1) is a MASP-1/3-associated plasma collectin with microbial-binding activity. J. Immunol. 185, 6096-6104.

Henriksen, M.L., Brandt, J., Andrieu, J.P, Nielsen, C., Jensen, P.H., Holmskov, U., Jorgensen, T.J., Palarasah, Y., Thielens, N.M., Hansen, S., 2013. Heteromeric complexes of native collectin

kidney 1 and collectin liver 1 are found in the circulation with MASPs and activate the complement system. J. Immunol. 191, 6117-6127.

Henriksen, M.L., Madsen, K.L., Skjoedt, K., Hansen, S., 2014. Calcium-sensitive immunoaffinity chromatography: Gentle and highly specific retrieval of a scarce plasma antigen, collectin-LK (CL-LK). J. Immunol. Methods 413, 25-31.

Hirao, T., Y, Sakamoto., M, Kamada., S, Hamada., T, Aono., 1993. Tn antigen, a marker of potential for metastasis of uterine cervix cancer cells. Cancer 72, 154-159.

Hirani, S., Lambris, J.D., Müller-Eberhard, H.J., 1985. Localization of the conglutinin binding site on the third component of human complement. J. Immunol. 134, 1105-1109.

Hogenkamp, A., Isohadouten, N., Reemers, S.S., Romijn, R.A, Hemrika, W., White, M.R., Tefsen, B., Vervelde, L., van Eijk, M., Veldhuizen, E.J., Haagsman, H.P., 2008. Chicken lung lectin is a functional C-type lectin and inhibits haemagglutination by influenza A virus. Vet, Microbiol. 130, 37-46.

Holmskov, U., Jensenius, J.C., 1993. Structure and function of collectins: humoral C-type lectins with collagenous regions. Behring Inst. Mitt. 93, 224-235.

Jang, S., Ohtani, K., Fukuoh, A., Yoshizaki, T., Fukuda, M., Motomura, W., Mori, K., Fukuzawa, J., Kitamoto, N., Yoshida, I., Suzuki, Y., Wakamiya, N., 2009. Scavenger receptor collectin placenta 1 (CL-P1) predominantly mediates zymosan phagocytosis by human vascular endothelial cells. J. Biol. Chem. 284, 3956–3965.

Jang, S., Ohtani, K., Fukuoh, A., Mori, K., Yoshizaki, T., Kitamoto, N., Kim, Y., Suzuki, Y., Wakamiya, N., 2014. Scavenger receptor CL-P1 mediates endocytosis by associating with AP-2μ2. Biochim. Biophys. Acta 1840, 3226-3237.

Keshi, H., Sakamoto, T., Kawai, T., Ohtani, K., Katoh, T., Jang, S.J., Motomura, W., Yoshizaki, T., Fukuda, M., Koyama, S., Fukuzawa, J., Fukuoh, A., Yoshida, I., Suzuki, Y., Wakamiya, N., 2006. Identification and characterization of a novel human collectin CL-K1. Microbiol. Immunol. 50, 1001–1013.

Kjaer, T., Jensen, L., Hansen, A., Dani, R., Jensenius J.C., Dobó, J., Gál. P., Thiel, S., 2016. Oligomerization of mannan-binding lectin dictates binding properties and complement activation. Scand. J. Immunol. *In press.*

Kolarova, J., Tangen, I., Bens, S., Gillessen-Kaesbach, G., Gutwein, J., Kautza, M., Rydzanicz, M., Stephani, U., Siebert, R., Ammerpohl, O., Caliebe, A., 2015. Array-based DNA methylation

analysis in individuals with developmental delay/intellectual disability and normal molecular karyotype. Eur. J. Med. Genet. 58, 419-425.

Koyama, S., Ohtani, K., Fukuzawa, J., Yao, N., Fukuda, M., Jang, S.J., Hasebe, N., Kikuchi, K., Itabe, H., Yoshida, I., Suzuki, Y., Wakamiya, N., 2011. The induction of human CL-P1 expression in hypoxia/reoxygenation culture condition and rat CL-P1 after ischemic/reperfusion treatment. Biochim. Biophys. Acta 1810, 836–842.

Ma, Y. J., Skjoedt, M.O., Garred, P., 2013. Collectin-11/MASP complex formation triggers activation of the lectin complement pathway–the fifth lectin pathway initiation complex. J. Innate. Immun. 5, 242–250.

Ma, Y.J., Hein, E., Munthe-Fog, L., Skjoedt, M.O., Bayarri-Olmos, R., Romani, L, Garred, P., 2015. Soluble collectin-12 (CL-12) is a pattern recognition molecule initiating complement activation via the alternative pathway. J. Immunol. 195, 3365–3373.

Mori, K., Ohtani, K., Jang, S., Kim, Y., Hwang, I., Roy, N., Matsuda, Y., Suzuki, Y., Wakamiya, N., 2014. Scavenger receptor CL-P1 mainly utilizes a collagen-like domain to uptake microbes and modified LDL. Biochim. Biophys. Acta 1840, 3345-3356.

Motomura, W., Yoshizaki, T., Ohtani, K., Okumura, T., Fukuda, M., Fukuzawa, J., Mori, K., Jang, S.J., Nomura, N., Yoshida, I., Suzuki, Y., Kohgo, Y., Wakamiya, N., 2008. Immunolocalization of a novel collectin CL-K1 in murine tissues. J. Histochem. Cytochem. 56, 243-252.

Nakamura K, Funakoshi H, Miyamoto K, Tokunaga F, Nakamura T. 2001. Molecular cloning and functional characterization of a human scavenger receptor with C-type lectin (SRCL), a novel member of a scavenger receptor family. Biochem. Biophys. Res. Commun. 280, 1028-1035.

Nakamura, K., Ohya, W., Funakoshi, H., Sakaguchi, G., Kato, A., Takeda, M., Kudo, T., Nakamura, T., 2006. Possible role of scavenger receptor SRCL in the clearance of amyloid-β in Alzheimer's disease. J. Neurosci. Res. 84, 874–890.

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

Ohtani, K., Suzuki, Y., Eda, S., Kawai, T., Kase, T., Keshi, H., Sakai, Y., Fukuoh, A., Sakamoto, T., Itabe, H., Suzutani, T., Ogasawara, M., Yoshida, I., Wakamiya, N., 2001. The membrane-type collectin CL-P1 is a scavenger receptor on vascular endothelial cells. J. Biol. Chem. 276, 44222-44228.

Peng, D., Wang, J., Zhang, R., Jiang, F., Tang, S., Chen, M., Yan, J., Sun, X., Wang, S., Wang,T., Yan, D., Bao, Y., Hu, C., Jia, W., 2015. Common variants in or near ZNRF1, COLEC12, SCYL1BP1 and API5 are associated with diabetic retinopathy in Chinese patients with type 2 diabetes. Diabetologia 58, 1231-1238.

Reid, K.B., Porter, R. R., 1976. Subunit composition and structure of subcomponent C1q of the first component of human complement. Biochem. J. 155, 19-23.

Reid, K.B., 1993. Structure/function relationships in the collectins (mammalian lectins containing collagen-like regions). Biochem. Soc. Trans. 21, 464-468.

Rooryck, C., Diaz-Font, A., Osborn, D.P., Chabchoub, E., Hernandez-Hernandez, V., Shamseldin, H., Kenny, J., Waters, A., Jenkins, D., Kaissi, A.A., Leal, G.F, Dallapiccola, B., Carnevale, F., Bitner-Glindzicz, M., Lees, M., Hennekam, R., Stanier, P., Burns, A.J., Peeters, H., Alkuraya, F.S., Beales, P.L., 2011. Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. Nat. Genet. 43, 197-203.

Roy, N., Ohtani, K., Matsuda, Y., Mori, K., Hwang, I., Suzuki, Y., Inoue, N., Wakamiya, N., 2016. Collectin CL-P1 utilizes C-reactive protein for complement activation. Biochim. Biophys. Acta 1860, 1118-1128.

Sastry, K., Ezekowitz, R.A., 1993. Collectins: pattern recognition molecules involved in first line host defense. Curr. Opin. Immunol. 5, 59-66.

Selman, L., Skjodt, K., Nielsen, O., Floridon, C., Holmskov, U., Hansen, S., 2008. Expression and tissue localization of collectin placenta 1 (CL-P1, SRCL) in human tissues. Mol. Immunol. 45, 3278-3288.

Selman, L,. Hansen, S., 2012. Structure and function of collectin liver 1 (CL-L1) and collectin 11 (CL-11, CL-K1). Immuonobiology. 217, 851-863.

Springer, G.F., R. Cheingsong-Popov, V., Schirrmacher, P.R. Desai, H, Tegtmeyer., 1983. Proposed molecular basis of murine tumor cell-hepatocyte interaction. J. Biol. Chem. 258, 5702-5706.

Springer, G. F., 1984. T and Tn, general carcinoma autoantigens. Science 224, 1198-1206.

Srivastava, S., Chen, L., 2009. Comparison between the stochastic search variable selection and the least absolute shrinkage and selection operator for genome-wide association studies of rheumatoid arthritis. BMC Proc. 3, (Suppl 7) : S21.

Takahashi, K., Ohtani, K., Larvie, M., Moyo, P., Chigweshe, L., Van Cott, E. M., Wakamiya N. 2014. Elevated plasma CL-K1 level is associated with a risk of developing disseminated intravascular coagulation (DIC). J. Thromb. Thrombolysis 38, 331-338.

Troegeler, A., Lugo-Villarino, G., Hansen, S., Rasolofo, V., Henriksen, ML., Mori, K., Ohtani, K., Duval, C., Mercier, I., Bénard, A., Nigou, J., Hudrisier, D., Wakamiya, N., Neyrolles, O., 2015. Collectin CL-LK Is a Novel Soluble Pattern Recognition Receptor for Mycobacterium tuberculosis. PLoS ONE 10, e0132692.

Troldborg, A., Thiel, S., Jensen, L., Hansen, S., Laska, M.J, Deleuran, B, Jensenius, J.C, Stengaard-Pedersen, K., 2015. Collectin liver 1 and collectin kidney 1 and other complement-associated pattern recognition molecules in systemic lupus erythematosus. Clin. Exp. Immunol. 182, 132-138.

Venkatraman Girija, U., Furze, C.M., Gingras, A.R., Yoshizaki, T., Ohtani, K., Marshall, J.E., Wallis, A.K., Schwaeble, WJ., El-Mezgueldi, M., Mitchell, D.A., Moody, P.C., Wakamiya, N., Wallis, R., 2015. Molecular basis of sugar recognition by collectin-K1 and the effects of mutations associated with 3MC syndrome. BMC Biol. 13, 1-13.

von der Heyde, S., Wagner, S., Czerny, A., Nietert, M., Ludewig, F., Salinas-Riester, G., Arlt, D., Beißbarth, T., 2015. mRNA profiling reveals determinants of trastuzumab efficiency in HER2-positive breast cancer. Plos ONE 10, e0117818.

Yoshida, T., Tsuruta, Y., Iwasaki, M., Yamane, S., Ochi, T., Suzuki, R., 2003. SRCL/CL-P1 recognizes GalNAc and a carcinoma associated antigen, Tn antigen. J. Biochem. 133, 271–277.

Yoshizaki, T., Ohtani, K., Motomura, W., Jang, S.J., Mori, K., Kitamoto, N., Yoshida, I., Suzuki, Y., Wakamiya, N., 2012. Comparison of human blood concentrations of collectin kidney 1 and mannan-binding lectin. J. Biochem. 151, 57-64.

Figure legends.

Figure 1. Structure of CL-K1 and CL-LK. A) Composition of the heteromeric complex (CL-LK) found in the circulation made of CL-K1 and CL-L1. Orange polypeptide represents CL-K1. B) Structure of the CL-K1 neck-CRD fragment of a trimeric homomeric subunit of CL-K1 obtained by crystallography. C) Interactions between CL-K1 and the disaccharide D-mannose (a1-2)-D-mannose, obtained by crystallography. The non-reducing mannose moiety (terminal) is in white and the penultimate sugar moiety (second to terminal) is in grey. The Ca2+ is in pink and polar interactions are shown by dotted lines. Panel A and B were modified from the CL-K1 crystal structure published by Venkatraman et al. 2015.

Figure 2. The three activation pathways and effector functions of the complement system. CL-LK represents the heteromeric complex of CL-L1 and CL-K1, found in the circulation.

Figure 3. Structure of the CRD from mouse CL-P1 bound to Lewisx (A) and illustration of the interactions between CL-P1, Ca2+ and Lewisx in the binding site (calcium binding site 2) (B). The mouse CL-P1 CRD is shown in cyan and Lewisx in grey, with bound Ca2+ in green. Selected coordination bonds between the carbohydrate or protein side chains, and the Ca2+ are shown in black dashed lines. Selected hydrogen bonds between the protein and the carbohydrate are shown in gray dashed lines. Hydrophobic interactions between the sugar and the protein or within the protein are in blue dashed lines. The figure was modified from the crystal structures presented in Feinberg et al. 2007.

Figure 4. Two pathways of complement activation using CL-P1. First, CL-P1 traps CRP on the B face and C1q binds to the A-face and activates the complement system through the classical and properdin mediated amplification pathway. The binding of CRP to CL-P1 also recruits CFH, which reduces the downstream activation of the complement and might inhibit the assembly of terminal complement complex. Second, shedding of soluble form of CL-P1 from umbilical cord tissue results soluble form of CL-P1 in circulation, which can recognize microbes. soluble CL-P1 then recruits properdin, thereby allowing alternative pathway activation leading to TCC formation on target surfaces. The figure was modified from the figure presented in Roy et al. 2016.

Table

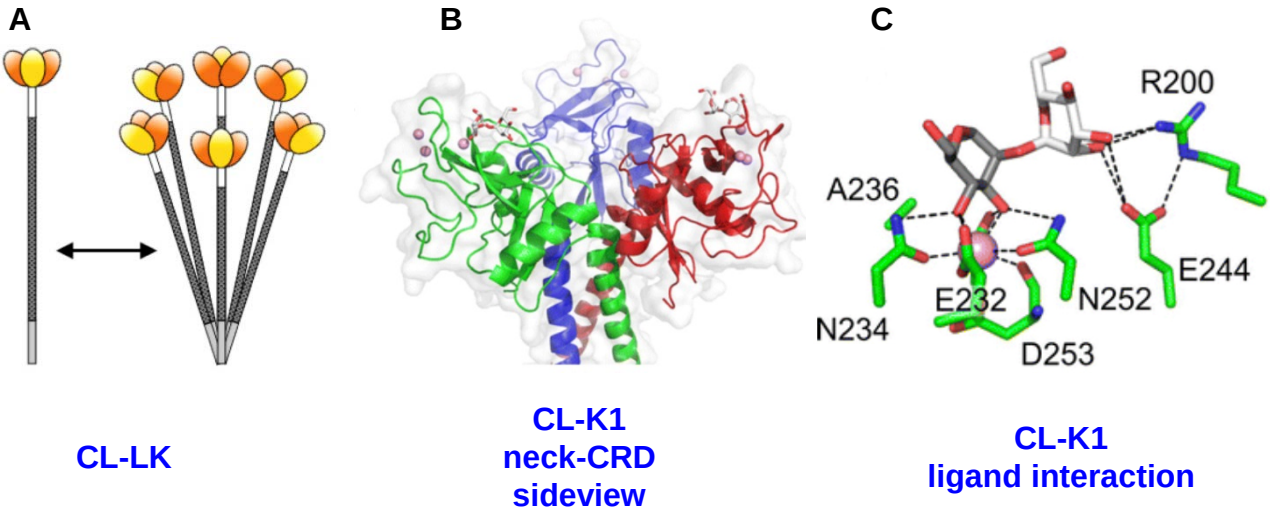
Table 1 **Characteristics and possible biological functions of human CL-L1, CL-K1 and CL-P1**

Collectins	Gene name	Other names	Tissue distribution	Sugar specificity	Ligand specificity	Possible Functions	Disease associations
CL-L1	*COLEC10*	collectin 10	liver, stomach, small intestine, colon, prostate, placenta, lung, heart, spleen, brain, kidney, bone marrow, mammary gland, uterus, pancreas, adrenal gland and thyroid	(D-Man, L-Fuc, D-Fuc, Gal, GlcNAc)*	n.d.	activation of lectin complement pathway in form of the complex CL-LK**	SLE
CL-K1	*COLEC11*	collectin 11	adrenals, liver, kidney, pancreas, CNS, stomach, small intestine, colon, testis, ovaries, placenta, lung, heart, thymus, large intestine, skeletal muscle and white adipose tissue	Mannose(α 1-2)-Mannose, D-Man, L-Fuc, GlcNAc	DNA, RNA, sulphated compounds, apoptotic cells, *E.coli*, *Klebsiella pneumoniae*, *Pseudomonus aeruginosa*, *Mycobacterium tuberculosis*, *Saccharomyces cerevisiae*, *Candida albicans* and Influenza A virus	activation of lectin complement pathway in form of the complex CL-LK, innate immunity to microorganisms, apoptotic cell clearance	3MC syndrome, DIC, urinary schistosomiasis and tuberculosis
CL-P1	*COLEC12*	collectin 12, SRCL	placenta, brain, heart, kidney, lung, bone marrow, colon, small intestine, stomach, thymus, mammary gland, prostate, skeletal muscle, testis, uterus, cerebellum and retina	$Lewis^X$, Gal, T-antigen, Tn-antigen,	OxLDL, zymosan, CRP, properdin, *E.coli*, *Staphylococcus aureus*	regulation of modified LDL in the circulation, activation of classical and alternative complement pathway via CRP, innate immunity to microorganisms and developmental process	Alzheimer's disease, rheumatoid arthritis, breast and thyroid cancer and diabetic retinopathy

D-man, D-mannan; L-Fuc, L-fucose; D-Fuc, D-fucose; GlcNAc, N-acetyl-D-glucosamine; Gal, galactose; OxLDL, Oxidized LDL; CRP, C-reactive protein; SLE, systemic lupus erythematosus; DIC, disseminated intravascular coagulation; n.d., not determined

*Due to heteromeric complex formation between CL-L1 and CL-K1, the reported specificity of CL-L1 might be influenced by the specificity of CL-K1

** CL-LK refers to the heteromeric form of CL-L1 and CL-K1 found in the circulation



CL-LK

CL-K1
neck-CRD
sideview

CL-K1
ligand interaction

Figure 1

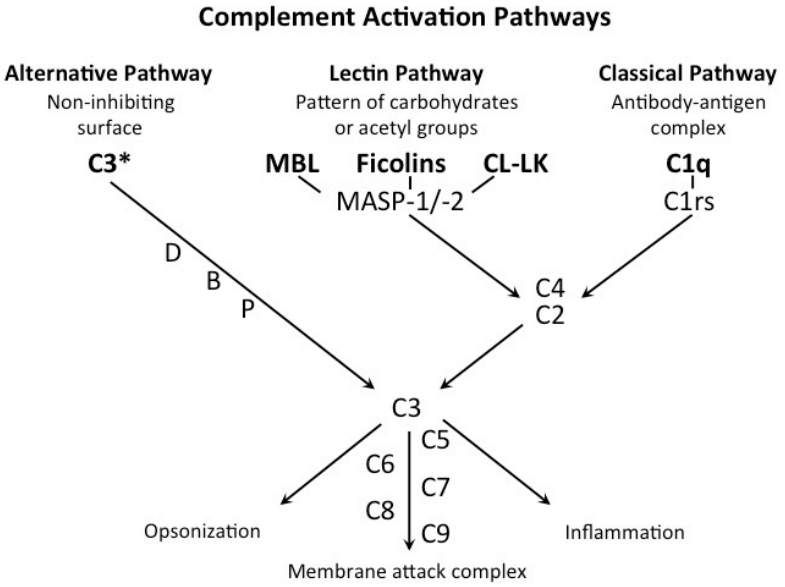

Complement Activation Pathways
Alternative Pathway
Non-inhibiting surface
C3*
Lectin Pathway
Pattern of carbohydrates or acetyl groups
MBL
Ficolins
CL-LK
MASP-1/-2
Classical Pathway
Antibody-antigen complex
C1q
C1rs
D
B
P
C4
C2
C3
C5
C6
C7
C8
C9
Opsonization
Inflammation
Membrane attack complex

Figure 2

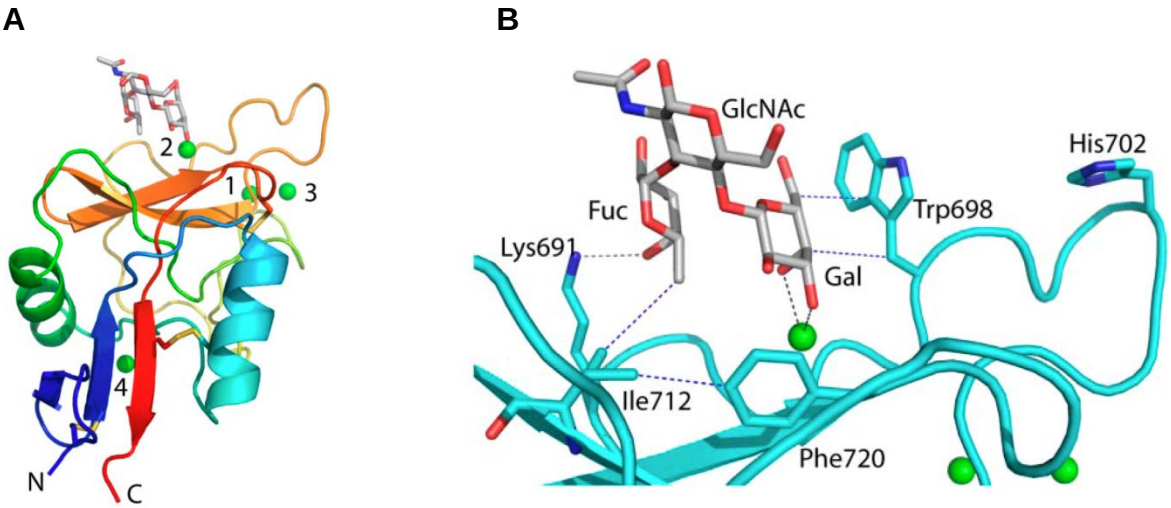

A
B
2
1
3
4
N
C
GlcNAc
His702
Fuc
Trp698
Lys691
Gal
Ile712
Phe720

Figure 3

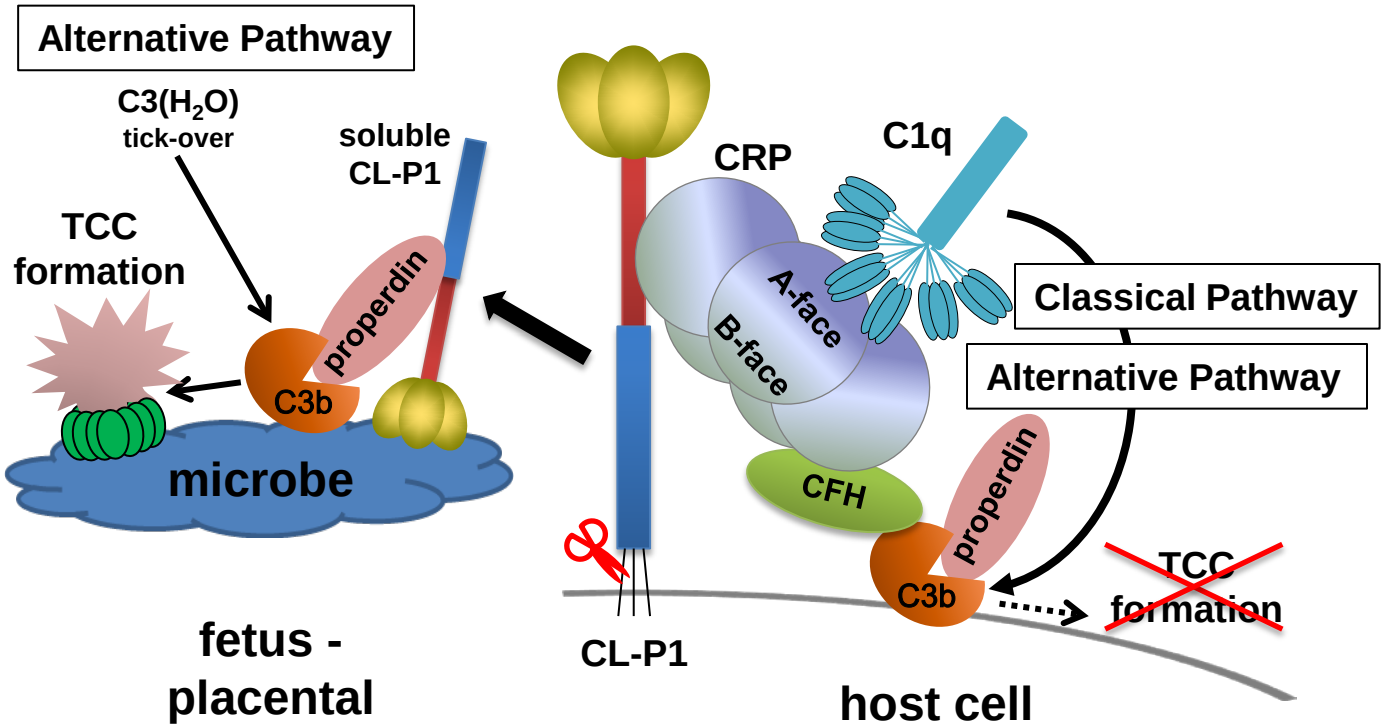

Alternative Pathway
C3(H2O)
tick-over
soluble
CL-P1
TCC
formation
properdin
C3b
microbe
fetus -
placental
CRP
C1q
A-face
B-face
Classical Pathway
Alternative Pathway
CFH
properdin
C3b
TCC
formation
CL-P1
host cell

Figure 4